FA-06 RESERVOIR NATIVE FISH GENETICS BASELINE STUDY REPORT

SKAGIT RIVER HYDROELECTRIC PROJECT FERC NO. 553

Seattle City Light

Prepared by:
Cramer Fish Sciences

March 2023
Updated Study Report

TABLE OF CONTENTS

Section No.	Description	Page No.
1.0	**Introduction**	**1-1**
2.0	**Study Goals and Objectives**	**2-1**
3.0	**Study Area**	**3-1**
4.0	**Methods**	**4-1**
4.1	Introduction and Background on Analysis Approach	4-1
4.1.1	Overview of Analyses	4-1
4.2	Data Requests and Sample Collections	4-10
4.2.1	Year 1 Data Requests and Sample Collections	4-10
4.2.2	Year 2 Data Requests and Sample Collections	4-11
4.3	Year 1 Genetic Analysis	4-14
4.3.1	Rainbow Trout	4-14
4.3.2	Bull Trout	4-15
4.3.3	Lineage Relationships	4-16
4.4	Year 2 Genetic Analysis	4-16
4.4.1	Year 2 Laboratory methods	4-16
4.4.2	Year 2 Hybridization	4-17
4.4.3	Year 2 Genetic Diversity within Collections	4-18
4.4.4	Year 2 Genetic Divergence among Collections	4-19
4.4.4.1	Pairwise F_{ST} between collections	4-19
4.4.4.2	Discriminant Analysis of Principal Components	4-19
4.4.4.3	Year 2 STRUCTURE Analysis	4-19
4.4.5	Year 2 Effective Population Size	4-20
4.4.6	Year 2 Age Determination for Effective Size Estimation	4-20
4.4.6.1	Scale Analysis	4-20
4.4.6.2	Length-at-Age Key	4-20
4.4.7	Year 2 Haplotype Diversity Associated with Life-history	4-21
5.0	**Results**	**5-1**
5.1	Expert Panel	5-1
5.1.1	Expert Panel Meeting No. 1	5-1
5.1.2	Expert Panel Meeting No. 2	5-2
5.1.3	Expert Panel Working Session	5-2
5.1.4	Expert Panel Coordination Meeting	5-2
5.1.5	Expert Panel Meeting No. 3	5-2
5.2	Year 1 Results	5-3
5.2.1	Year 1 Rainbow Trout (*Oncorhynchus mykiss*)	5-3

5.2.1.1 Year 1 Collections (existing data) 5-3
5.2.1.2 Year 1 Genetic Summary Statistics 5-3
5.2.2 Year 1 Native Char (*Salvelinus* spp.) 5-12
5.2.2.1 Year 1 Collections (existing data) 5-12
5.2.2.2 Year 1 Identification of Related Individuals within Collections 5-13
5.2.2.3 Year 1 Population Determination 5-15
5.2.2.4 Year 1 Genetic summary statistics 5-17
5.3 Year 2 Results 5-19
5.3.1 Year 2 Rainbow Trout (*Oncorhynchus mykiss*) 5-19
5.3.1.1 Year 2 Hybridization 5-19
5.3.1.2 Year 2 Within and Among Population Diversity 5-21
5.3.1.3 Year 2 Scale Age Determination 5-32
5.3.1.4 Length-at-Age Key Assignment 5-32
5.3.1.5 Year 2 Effective Population Size 5-33
5.3.1.6 Year 2 Haplotype Diversity 5-34
5.3.1.7 Year 2 Above and Below Project Population Analysis 5-35
5.3.1.8 Year 2 Regional Population Analysis 5-39
5.3.2 Year 2 Native Char (*Salvelinus* spp.) 5-42
5.3.2.1 Year 2 Hybridization 5-42
5.3.2.2 Genetic variation within collections of Bull Trout and Dolly Varden 5-44
5.3.2.3 Year 2 Genetic divergence among collections of Bull Trout and Dolly Varden 5-47
5.3.2.4 Year 2 Discriminant Analysis of Principal Components 5-49
5.3.2.5 Year 2 Scale Age Determination for Estimating N_e and N_b 5-56
5.3.2.6 Year 2 Length-at-Age Key Assignment 5-57
5.3.2.7 Year 2 Effective population size 5-58
6.0 Discussion and Findings 6-1
6.1 Year 1 6-1
6.1.1 Summary of Completed Objectives 6-1
6.2 Year 2 6-2
6.2.1 Summary of Completed Objectives 6-2
6.2.2 Discussion of Rainbow Trout (*Oncorhynchus mykiss*) 6-3
6.2.2.1 Conclusions 6-4
6.2.3 Discussion of Native Char (*Salvelinus* spp.) 6-5
6.2.3.1 Conclusion 6-7
6.3 Status of June 9, 2021 Notice 6-8

7.0 Variances from Proposed Study Plan and Proposed Modifications 7-1

8.0 References 8-1

List of Figures

Figure No.	Description	Page No.
Figure 3.0-1.	Map showing the various collection locations of *O. mykiss* analyzed in year 1. Year 1 data were provided in an ad hoc fashion and so important information about the specific locations of collections was not available. Some collections were made within the study area (outlined in red), while others were outside of it. For samples collected downstream of Gorge Dam, it is not known whether collections were made upstream or downstream of barriers or within the anadromous zone, because the metadata regarding the sampling locations was not provided. See Table 5.2-1 for details on anadromous barrier locations identified in Year 2 field surveys.	3-1
Figure 3.0-2.	Map showing the collection locations of Bull Trout analyzed in Year 1. Year 1 data were provided in an ad hoc fashion and so important information about the specific locations of collections was not available. Some collections were made within the study area (outlined in red), while others were outside of it. For samples collected downstream of Gorge Dam, it is not known whether collections were made upstream or downstream of barriers or within the anadromous zone, because the metadata regarding the sampling locations was not provided. The USFWS samples were assumed to be at the same locations as reported in Smith (2010), but no location data was provided. Samples obtained from within the study area were considered "at-large" from reservoirs, as no location information was available from the WDFW other than the collection code. See Table 5.2-1 for details on anadromous barrier locations identified in Year 2 field surveys.	3-2
Figure 4.2-1.	Year 2 *Salvelinus* and *Oncorhynchus* genetics sampling locations. Primary tributaries were identified by direct connectivity to the reservoirs and divided into three sampling strata, a low, mid, and high elevation, to avoid overrepresentation of the family groups. Secondary tributaries are identified by connectivity to a primary tributary, do not flow directly into the reservoirs, and were only sampled at a single location. Locations such as Silver mid and high, Thunder high, Colonial mid and high, were not sampled due to challenges during the field season. Other locations, such as Luna Creek and Hozomeen Creek, were sampled but no fish were collected.	4-13
Figure 5.2-1.	Log-likelihood ratio distribution for simulated true full-siblings versus unrelated individuals based on Roland Creek *O. mykiss* genotype data from Year 1. The analysis used genotypes from 15 microsatellites that were provided by WDFW. Legend: FS=Full Sibling; U=Unrelated. High overlap between full-siblings and unrelated fish suggests relatively low statistical power to detect highly related individuals.	5-4

Figure 5.2-2. Scatterplot of PC-1 (2.215 percent) and PC-2 (2.044 percent) for Year 1 Rainbow Trout microsatellite data for all collections, excluding samples from the Baker River, estimated using {adegenet} in program R (Jombart 2011). The distribution of genetic variation appears to support some genetic structuring associated with the geographical locations of collections. River basins are provided to show the approximate geographical locations of each collection. NF Cascade was apparently collected upstream of a barrier. Metadata and types of collections (i.e., resident/adult) are shown in Table 5.2-1. Numbers at centroids identify the collection number listed in Table 5.2-1. Ellipses define 1.5 standard deviations of the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom left corner shows first three eigenvalues. 5-9

Figure 5.2-3. Scatterplot of genetic PC-1 (2.215 percent) and PC-3 (1.836 percent) for all Rainbow Trout collections, excluding samples from the Baker River, estimated using {adegenet} in program R. The distribution of genetic variation appears to support existence of genetic structuring associated with the geographical locations of collections. River basin names are provided to describe the approximate geographical locations of each collection. Numbers at centroids identify the collection number listed in Table 5.2-1. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom right corner shows first three eigenvalues. 5-10

Figure 5.2-4. Scatterplot of genetic PC-1 (explaining 3.870 percent of the variation) and PC-2 (explaining 2.028 percent of the variation) for all Rainbow Trout collections located upstream of the Gorge Dam. The plot was generated using the {adegenet} package in the R programming language and 15 microsatellites from Year 1 were used. The numbers at the centroids correspond to the collection number listed in Table 5.2-1. The scree plot in the bottom right corner displays the first three eigenvalues. The ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more indicate less distinction between groups. The scree plot in the bottom left corner also shows the first three eigenvalues. It is important to note that inertia is a measure of relative genetic differences between individuals (as outlined in the Preamble of the Methods section), and should not be confused with F_{ST}, which measures divergence at the population level. 5-11

Figure 5.2-5. Log likelihood ratios distribution for simulated true full siblings versus unrelated individuals based on Skagit River *S. confluentus* genotype Year 1 microsatellite data provided by WDFW and USFWS (Table 5.2-2). High overlap between full-siblings and unrelated fish suggests relatively low power to detect highly related individuals using the 16 microsatellites. Legend: FS=Full Sibling; U=Unrelated. High overlap between full-siblings and unrelated fish suggests relatively low statistical power to detect highly related individuals. 5-14

Figure 5.2-6. Visualization of k-means clustering analysis at k=5 for Bull Trout individuals from previously reported microsatellite dataset at 1st and 2nd

principal component axes. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in upper right corner shows first three eigenvalues. Cluster 1 were study area tributary collections and contained a majority of Ross Lake samples. Cluster 2 were individuals collected from below Gorge Dam. Clusters 3, 4 and 5 were predominantly individuals collected from Diablo and Gorge lakes (see Figure 5.2-8 for refined locations). 5-16

Figure 5.2-7. Genetic clusters visualized in Figure 5.2-7 aligned to each Bull Trout collection from previously reported microsatellite dataset. Size of boxes is scaled by sample count. Genetic clusters are organized by geographic location with upper Skagit collections at the top and lower Skagit at the bottom. Inferred clusters (inf 1-5) are the same as shown in Figure 5.2-7. 5-17

Figure 5.3-1. Proportional distribution of *O. mykiss* hybridization index. Non-hybrid *O. mykiss* make up 87 percent of collected samples, while the remaining 13 percent of hybridized individuals observed, ranged from 1-6 on the hybrid index. Proportions are conditional on resolution provided by diagnostic SNP loci present in SNP panel. 5-19

Figure 5.3-2. Year 2 map showing the distribution of the proportion of hybridized *Oncorhynchus* individuals collected across the Project area. 5-20

Figure 5.3-3. Visual representation of DAPC analysis on 1st and 2nd principal components axes of Study Area *O. mykiss* for *k*=4 genetic clusters. Cluster 1 contained a majority of Study Area *O. mykiss*, cluster 2 was Little Beaver Creek, cluster 3 was mostly Three Fools Creek, and cluster 4 was Pyramid Creek. 5-25

Figure 5.3-4. Visual representation of DAPC analysis on 1st and 2nd principal components axes of Project *O. mykiss* for *k*=5 genetic clusters. Cluster 1 was identified as Little Beaver Creek, cluster 2 as Pyramid Creek, cluster 4 as majority Three Fools Creek. Individuals assigning to Clusters 3 and 5 were widely distributed across remaining tributaries in Study Area. 5-26

Figure 5.3-5. Visual representation of membership probabilities given the same *O. mykiss* k=5 genetic clusters shown in Figure 5.3-4. Individual fish are displayed approximate north to south (Silver Creek starting on the right and ending with Gorge Lake on left). Cluster 1 was Little Beaver Creek, cluster 2 was Pyramid Creek, and cluster 4 was majority Three Fools Creek. 5-27

Figure 5.3-6. Visual representation of HWE estimates, with locus by population combination failing test at α less than 0.05 shown in pink. In this data configuration, a majority of loci conformed to HWE expectations, and no locus failed HWE across all population. 5-28

Figure 5.3-7. Pairwise relatedness (Rxy) between individual *O. mykiss* within each Project population containing greater than 15 samples. For all populations except Lightning and Three Fools Creeks, the mean Rxy was below zero. An Rxy = 0.0, 0.25, and 0.50 equates to individuals being unrelated, half siblings, and full siblings, respectively. 5-29

Figure 5.3-8. Genetic diversity (mean observed and mean expected heterozygosity) for Project *O. mykiss* populations containing greater than 15 samples. 5-30

Figure 5.3-9. Pairwise estimates of F_{ST} for Study Area *O. mykiss* populations containing greater than 15 samples. All pairwise F_{ST} estimates were statistically significant except for Granite-3 versus Canyon-3. F_{ST} values shown are the actual F_{ST} quantities, and not the significance of each pairwise test. 5-31
Figure 5.3-10. Summary of age, based on scale analysis, and fork length (mm) of 84 *Oncorhynchus* captured in 2022. 5-32
Figure 5.3-11. Summary of age, based on scale analysis and age-assignment, and fork length (mm) of 508 *Oncorhynchus*. 5-33
Figure 5.3-12. Frequency of the most common OMY5 haplotype (haplotype 3) in Project *O. mykiss* populations. 5-34
Figure 5.3-13. Visual representation of DAPC analysis on 1st and 2nd principal components axes of above- and below-Project *O. mykiss* for k=5 genetic clusters. First principal component (x axis) pertains to above and below Gorge Dam variance, while 2nd principal component (y axis) was driven by differences at Little Beaver Creek. Cluster 1 was predominantly Three Fools Creek with some Lightning Creek individuals included. Cluster 2 consisted of most of the Project *O. mykiss* populations. Cluster 3 was primarily Pyramid Creek. Individuals from below Project populations resided in cluster 4, with the exception of one individual, and cluster 5 was Little Beaver Creek. 5-36
Figure 5.3-14. Genetic diversity (mean observed and mean expected heterozygosity) for above- and below-Project *O. mykiss* populations containing greater than 15 samples. 5-37
Figure 5.3-15. Pairwise estimates of F_{ST} for above- and below-Gorge Dam O. mykiss populations containing greater than 15 samples. F_{ST} values shown are the actual F_{ST} quantities, and not the significance of each pairwise test. 5-38
Figure 5.3-16. Visual representation of DAPC analysis on 1st and 2nd principal components axes of regional *O. mykiss* dataset for k=5 genetic clusters. First principal component (x axis) pertains to differences between *O. mykiss* subspecies (coastal versus interior redband), while 2nd principal component (y axis) was driven by differences of Project Area *O. mykiss*. Cluster 3 represents 29 (of 30) Project *O. mykiss* populations. The Pyramid Creek population resides in Cluster 1 with other coastal *O. mykiss* populations (*O. m. irideus*). Clusters 2, 4, and 5 represent inland redband (*O. m. gairdneri*) populations in regional dataset. 5-41
Figure 5.3-17. Map of Year 2 *Salvelinus* collections showing the distribution of the proportion of individuals that were Bull Trout, Dolly Varden, Brook Trout, or hybrids across the Project area based on 22 taxon-diagnostic SNPs. 5-43
Figure 5.3-18. Scatterplot of first 2 PCs based on genotypes at 263 GTseq SNPs within *Salvelinus* collected in the study area during Year 2 (summer/fall 2022). 5-44
Figure 5.3-19. Isolation by distance analyses for Dolly Varden in the study area assayed at 8 microsatellite loci. Linear pairwise F_{ST} distances are plotted against pairwise geographical distance. The Mantel test suggests that 40 percent of the variability observed in the F_{ST} is explained by geographic distance. 5-48

Figure 5.3-20. Scatterplot of the squared residuals from the isolation-by-distance analysis in Dolly Varden using 8 microsatellite loci. Although a positive relationship was observed (R^2=0.10), the Mantel test was not statistically significant (P=0.11). The analysis was rerun without the apparent outlier (Point in the box, Colonial and Lightning) and the interpretation did not change (i.e., the Mantel test remained nonsignificant)..... 5-48

Figure 5.3-21. Scatterplot of the first 2 PCs based on 33 GT-Seq SNP genotypes in 65 Bull Trout sampled in Year 1 (2022) Diablo (centroid 1) Gorge (centroid 2) Ross (centroid 3) lakes. From the initial suite of 235 neutral markers, 200 were removed from the analysis due to lack of polymorphism (i.e., Minor Allele Frequency less than 0.01). Bull Trout were grouped by Reservoir because sample sizes were too small to analyze by tributaries. 5-50

Figure 5.3-22. Scatter plot of the first two linear discriminants produced by the final DAPC for 65 Bull Trout that were sampled during the 2022 field season (year 2) and genotyped using 33 GT-seq SNP markers. The analysis inferred three genetic clusters, which are identified by the numbers (in black) in each ellipse. The color and shape of each point, as shown in the legend, indicates whether the Bull Trout was sampled in Diablo, Gorge, or Ross Reservoirs. Genetic cluster 1 contained Bull Trout from all three reservoirs. Nevertheless, Bull Trout from Ross tended to have higher loadings for the first linear discriminant, while those from Gorge tended to have lower loadings. All three Bull Trout from Diablo were grouped within cluster 1, although this sample size was very small. The analysis used k=3, 6 Principal Components (PCs), and two discriminant functions. The specific driver of the genetic structure observed is unclear, but 100 percent accuracy of posterior assignments back to each cluster suggests genetic structure among Bull Trout is present..... 5-51

Figure 5.3-23. Composition plot that displays the posterior probability of assignment of 65 Bull Trout samples from the Project area to k=3 inferred DAPC clusters during year 2. Each vertical line in the plot represents an individual fish, and the color of the line represents its posterior probability of assignment to three inferred genetic clusters. The results indicate that Bull Trout did not cluster entirely by reservoir. The individuals are sorted based on their posterior probability of assignment to each of the three inferred clusters identified using DAPC. 5-53

Figure 5.3-24. Scatterplot of first 2 PCs based on genotypes of 413 Dolly Varden sampled in 13 collections during year 2. Analysis is based on genotypes at eight microsatellites. Each point represents an individual Dolly Varden. Note: (1) Big Beaver, (2) Canyon, (3) Colonial, (4) Granite, (5) Hozomeen, (6) Lightning, (7) NF Canyon, (8) Pierce, (9) Roland, (10) Ruby, (11) Silver, (12) Stetattle, (13) Thunder creeks. 5-54

Figure 5.3-25. Page 1 of 2. The scatterplot shows the discriminant function scores for seven microsatellites that were genotyped in Dolly Varden. The analysis assumes K=12 and uses 40 principal components. Each individual is represented by a point in the scatterplot. The inset in this panel displays the eigenvalues of

the analysis. The plot displays the projection of the first two linear discriminants. The x-axis represents the first linear discriminant, and the y-axis represents the second linear discriminant. Two panels are presented to facilitate interpretation of the results. In this panel, elipses are associated with the 12 inferred clusters. In the next panel, individuals are colored and shaped according to their tributary of origin. Both panels show the exact same information the points are just colored/shaped differently depending on whether they are displaying the inferred clusters (top panel) or tributary of origin (bottom panel) 5-55

Figure 5.3-25. Page 2 of 2. In the previous page, the scatterplot shows the discriminant function scores for seven microsatellites that were genotyped in Dolly Varden. The analysis assumes K=12 and uses 40 principal components. Each individual is represented by a point in the scatterplot. The inset in the previous panel displays the eigenvalues of the analysis. The plot displays the projection of the first two linear discriminants. The x-axis represents the first linear discriminant, and the y-axis represents the second linear discriminant. Two panels are presented to facilitate interpretation of the results. In the previous panel, elipses are associated with the 12 inferred clusters. In this panel, individuals are colored and shaped according to their tributary of origin. Both panels show the exact same information the points are just colored/shaped differently depending on whether they are displaying the inferred clusters (top panel) or tributary of origin (bottom panel)..... 5-56

Figure 5.3-26. Summary of age, based on scale analysis, and fork length (mm) of 110 *Salvelinus* captured in 2022. 5-57

Figure 5.3-27. Summary of age, based on scale analysis and age-assignment, and fork length (mm) of 311 *Salvelinus*. 5-58

List of Tables

Table No.	Description	Page No.
Table 4.1-1.	Glossary of terms, statistical test, test metric, purpose of test, significant value threshold, species tested.	4-7
Table 4.2-1.	Summary of locations[1] targeted for sampling *Oncorhynchus* and *Salvelinus* during the Year 2 field season in 2022.	4-12
Table 4.4-1.	The 22 diagnostic SNP markers for taxa within *Salvelinus* (Bohling et al. 2021). These markers were used alongside field samples to distinguish among Bull Trout, Brook Trout, and Dolly Varden collections from the Project area.	4-18
Table 5.2-1.	Year 1 summary statistics for samples collected from *O. mykiss* in the Skagit and Fraser River basins. Analysis of these data have been previously reported, for example, in Pflug et al. (2013). Genotypes consisted of 15 microsatellite loci. The collections were provided by WDFW. The collections came from various places, including the Project area,	

downstream of the Project area, and adjacent watersheds. Some collections were from the anadromous zone (identified in the table notes). 5-5
Table 5.2-2 Summary of Year 1 Bull Trout microsatellite dataset collection provided by WDFW (335) and USFWS (563). These samples were analyzed using 16 microsatellites and 589 were retained and evaluated. Suspected hybrids were removed before analysis. 5-13
Table 5.2-3. Year 1 summary statistics for samples collected from Bull Trout in the Skagit River basin. 5-18
Table 5.2-4. Table of pairwise estimates of F_{ST} between the Project area collections of Bull Trout. 5-19
Table 5.3-1. Non-hybrid *O. mykiss* samples used for Year 2 genetic analysis. 5-21
Table 5.3-2. Non-hybrid *O. mykiss* samples used for genetic analysis. Bold F_{IS} values were statistically significant from zero. 5-23
Table 5.3-3. Size and age summaries for scale-determined ages of *Oncorhynchus* individuals. 5-32
Table 5.3-4. Size and age summaries for aged and age-assigned *Oncorhynchus* individuals. 5-33
Table 5.3-5. Diversity of OMY5 haplotypes for Project *O. mykiss* individuals. 5-34
Table 5.3-6. Counts of regional DAPC cluster membership for Skagit Basin *O. mykiss* samples used for regional genetic analysis. Population labels were retained from Table 5.3-2. Note that no Skagit Basin individuals analyzed assigned to Clusters 2 and 5 (inland redband). 5-40
Table 5.3-7. Summary statistics[1] of 2022 Bull Trout (N=65) collections. 5-45
Table 5.3-8. Summary statistics for collections of Dolly Varden (N=413) and for the 12 inferred genetic groupings identified by DAPC (N=405). Only individuals with greater than 0.50 probability of assignment to an inferred cluster were included (N=405). Note, the twelve inferred clusters are depicted graphically in Figure 5.3-25. 5-46
Table 5.3-9. Pairwise F_{ST} for Bull Trout collection pools based on reservoirs and inferred genetic clusters from DAPC analysis. Statistically significance estimates (at alpha=0.05 level) are indicated by bold lettering. 5-49
Table 5.3-10. Size and age summaries for scale age determined *Salvelinus* individuals 5-57
Table 5.3-11. Size and age summaries for aged and age-assigned *Salvelinus* individuals. 5-57
Table 5.3-12. Effective population size estimates for Bull Trout corrected for overlapping generations using Waples et al. (2014) adjustment based on adult life span (8.5 years) and age of first reproduction (3.0) (Hemmingsen et al. 2001). Bias corrections are insensitive to Adult Lifespan and Age of Maturity within a few years. Only collections with greater than 20 samples were analyzed. 5-59
Table 5.3-13. Effective population size estimates for Dolly Varden corrected for overlapping generations using Waples et al. (2014) adjustment based on adult life span and age of first reproduction. 5-59

Table 6.3-1. Status of Stranding and Trapping Assessment modifications identified in the June 9, 2021 Notice. 6-8

List of Attachments

Attachment A Year 1 Existing Genetics Data Review Technical Memorandum
Attachment B Year 2 Sampling Plan
Attachment C Year 2 Technical Memorandum
Attachment D Expert Panel Meeting Summaries

AMOVA .....analysis of molecular variance
AR.....allelic richness
BIC.....Bayesian Information Criterion
CI.....confidence interval
City Light.....Seattle City Light
CFS .....Cramer Fish Sciences
CRITFC.....Columbia River Inter-Tribal Fish Commission
CKMR.....close kin mark recapture
DAPC.....discriminant analysis of principal components
F_{IS}.....inbreeding coefficient
F_{ST} .....genetic divergence
FERC.....Federal Energy Regulatory Commission
FNR.....false negative rate
FPR .....false positive rate
FWER .....family-wise error rate
GTseq.....Genotyping-in-Thousands by Sequencing
H_O.....observed multilocus heterozygosity within populations
H_S .....expected multilocus heterozygosity within populations
HOR .....Hatchery origin
HWE .....Hardy-Weinberg equilibrium
HWP.....Hardy-Weinberg Proportions
ILP.....Integrated Licensing Process
ISR .....Initial Study Report
K.....assumed population number
LD .....linkage disequilibrium
LOD .....log-of-the-odds
LP.....licensing participant
m .....Migration rate
mm .....millimeters
MNA .....mean number of alleles
N_b .....effective number of breeders

N_e .....effective population size
Nm.....migrants per generation
NMFS.....National Marine Fisheries Service
NOR .....Natural origin
PC.....principal component
PCA.....principal component analysis
PCR.....polymerase chain reaction
Project .....Skagit River Hydroelectric Project
RSP .....Revised Study Plan
SNP .....Single Nucleotide Polymorphism
SPAN .....Stevan Phelps Allele Nomenclature
µL.....micro liters
USFWS .....U.S. Fish and Wildlife Service
USGS .....U.S. Geological Survey
USR.....Updated Study Plan
WDFW.....Washington Department of Fish and Wildlife
YOY .....young-of-the-year

This page intentionally left blank.

1.0 INTRODUCTION

The FA-06 Reservoir Native Fish Genetics Baseline Study (Reservoir Fish Genetics Study) is being conducted in support of the relicensing of the Skagit River Hydroelectric Project (Project), Federal Energy Regulatory Commission (FERC) No. 553, as identified in the Revised Study Plan (RSP) submitted by Seattle City Light (City Light) on April 7, 2021 (City Light 2021). On June 9, 2021, City Light filed a "Notice of Certain Agreements on Study Plans for the Skagit Relicensing" (June 9, 2021 Notice)[1] that detailed additional modifications to the RSP agreed to between City Light and supporting licensing participants (LP) (which include the Swinomish Indian Tribal Community, Upper Skagit Indian Tribe, National Marine Fisheries Service [NMFS], National Park Service, U.S. Fish and Wildlife Service [USFWS], Washington State Department of Ecology, and Washington Department of Fish and Wildlife [WDFW]). The June 9, 2021 Notice included agreed to modifications to the Reservoir Fish Genetics Study.

In its July 16, 2021 Study Plan Determination, FERC did not require implementation of the Reservoir Fish Genetics Study. Notwithstanding, City Light implemented the Reservoir Fish Genetics Study as proposed in the RSP with the agreed to modifications described in the June 9, 2021 Notice.

On March 8, 2022, City Light filed its Initial Study Report (ISR). No requests for modifications to the Reservoir Fish Genetics Study were filed. FERC's August 8, 2022 Determination on Requests for Study Modifications required no modifications to the Reservoir Fish Genetics Study.

This study is complete and a report of the study efforts is being filed with FERC as part of City Light's Updated Study Report (USR).

[1] Referred to by FERC in its July 16, 2021 Study Plan Determination as the "updated RSP."

2.0 STUDY GOALS AND OBJECTIVES

The goals of this study are to characterize baseline population genetic structure for three native salmonid species: Bull Trout (*Salvelinus confluentus*), Rainbow Trout (*Oncorhynchus mykiss*), and Dolly Varden (*Salvelinus malma*) (target species) in Project reservoirs and provide the basis necessary to inform the planning of long-term (i.e., over the new license term) reservoir fish management objectives. Specifically, the goals of this study are to:

- Determine the population genetic structure within and among target species populations and assess whether management actions are necessary for genetic sustainability.
- Determine the number of fish populations, for each target species, within and among the Project reservoirs.
- Estimate the effective population size ($N_{e)}$ for each target species and reservoir.
- Identify topics and/or management objectives to be considered in the Project reservoir fish and aquatics management plan.

Specific objectives to meet these study goals are listed below.

Year 1

- City Light will convene an Expert Panel in consultation with LPs.
- Review, compile, and analyze target species genetics data collected by multiple researchers in the Project reservoirs.
- Acquire and consolidate existing genetics data for Bull Trout, Rainbow Trout, and Dolly Varden.
- Create a single, standardized data file for each species that compiles genotypes from existing studies.
- Use the standardized data files to evaluate baseline genetic metrics for Bull Trout and Rainbow Trout.
- Calculate within- and among-population summary statistics using consistent methods for Bull Trout and Rainbow Trout.
- Estimate relatedness for Bull Trout and Rainbow Trout and report the statistical distribution of this metric by species and reservoir.
- Estimate the power (false detection rate) of genetic markers currently in use to identify relationships (e.g., parent-offspring pairs, full-sibling-unrelated pairs).
- Identify the availability of relevant existing genetic samples and coordinate target fish species sampling being conducted opportunistically by other relicensing studies and current license field activities.
- Expert Panel review of Year 1 study results and assistance in development of Year 2 study program.

Year 2

- Expand sample collection and/or coordinate existing samples and activities for out-of-basin and above and below dam analyses.
- Continue data collection to address heterozygosity, within- and among-population variance, and relatedness for Dolly Varden in Project reservoirs.
- Gather additional data needed to estimate N_e for each population of Bull Trout, Rainbow Trout, and Dolly Varden.
- Gather the data needed to estimate N_e during the Integrated Licensing Process (ILP) study period.

Under the June 9, 2021 Notice, City Light and the supporting LPs agreed to four modifications to this Reservoir Fish Genetics Study:

- City Light will modify the study plan to collect juvenile fish at spawning grounds for genetics baseline as part of field sampling program in Year 2.
- City Light will modify the study plan to expand sample collection and/or coordination of existing samples and activities and analysis out-of-basin and above/below dams.
- City Light will clarify the study plan to explain the role of the Expert Panel. The LPs and City Light agree that: (1) the Expert Panel will serve in an advisory role; and (2) the Expert Panel will include experts from fields other than genetics.
- City Light will modify the study plan to provide that City Light will seek input from LPs and advice from an Expert Panel on whether and how genetics information or other monitoring methods can be used to inform future evaluation of reservoir fish abundance, habitat use, and migration timing.

3.0 STUDY AREA

The study area includes the Project reservoirs (i.e., Gorge, Diablo and Ross lakes in the U.S.) and associated reservoir tributaries (Figure 3.0-1). Additionally, because existing data is being used and consistent with the June 9, 2021 Notice, the geographic area of the Reservoir Fish Genetics Study was expanded to include sample collection/coordination of existing samples and activities, and analysis of out-of-basin areas and above/below the Project dams, including below Gorge Dam (Figures 3.0-1 and 3.0-2).

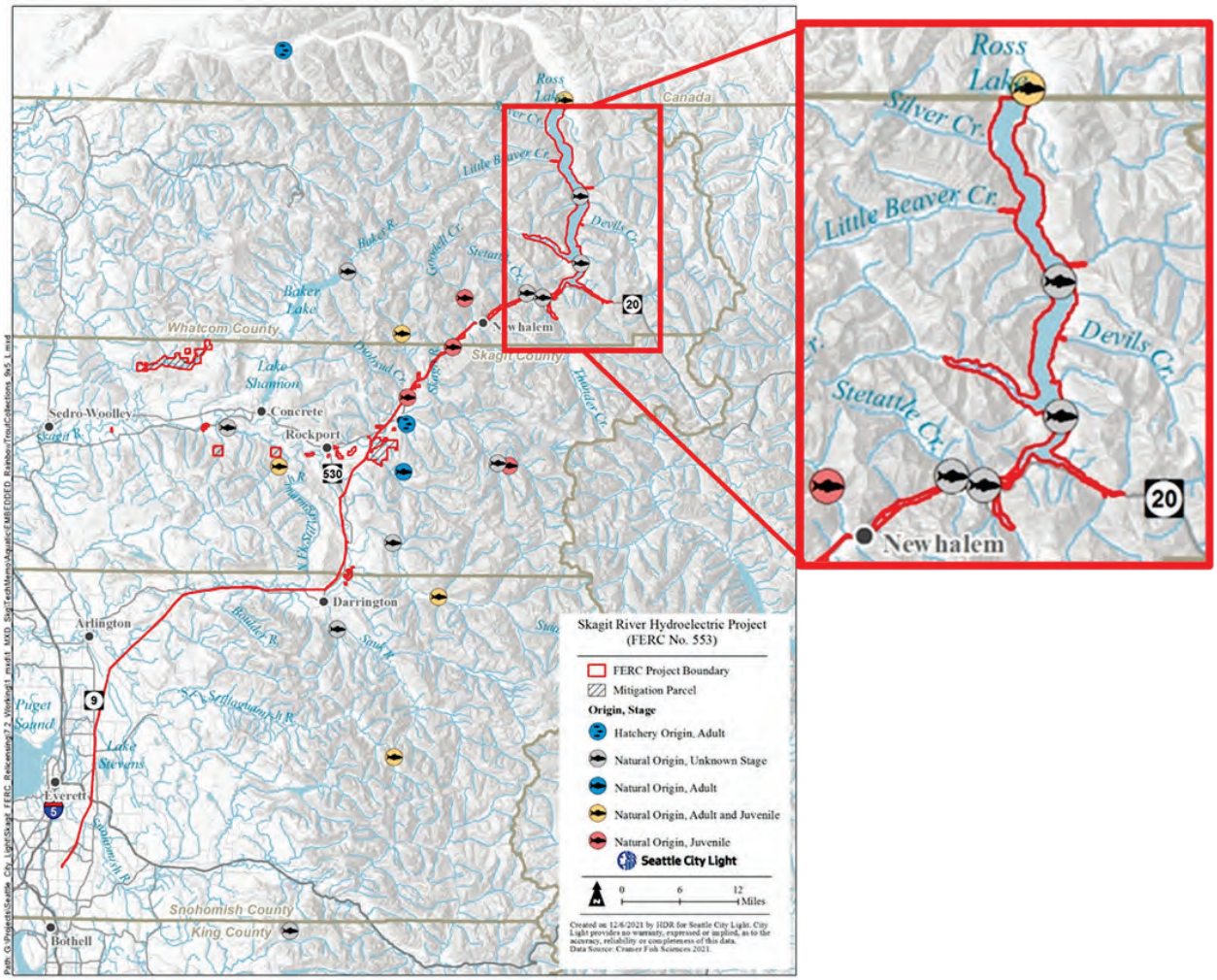


Figure 3.0-1. **Map showing the various collection locations of *O. mykiss* analyzed in year 1. Year 1 data were provided in an ad hoc fashion and so important information about the specific locations of collections was not available. Some collections were made within the study area (outlined in red), while others were outside of it. For samples collected downstream of Gorge Dam, it is not known whether collections were made upstream or downstream of barriers or within the anadromous zone, because the metadata regarding the sampling locations was not provided. See Table 5.2-1 for details on anadromous barrier locations identified in Year 2 field surveys.**

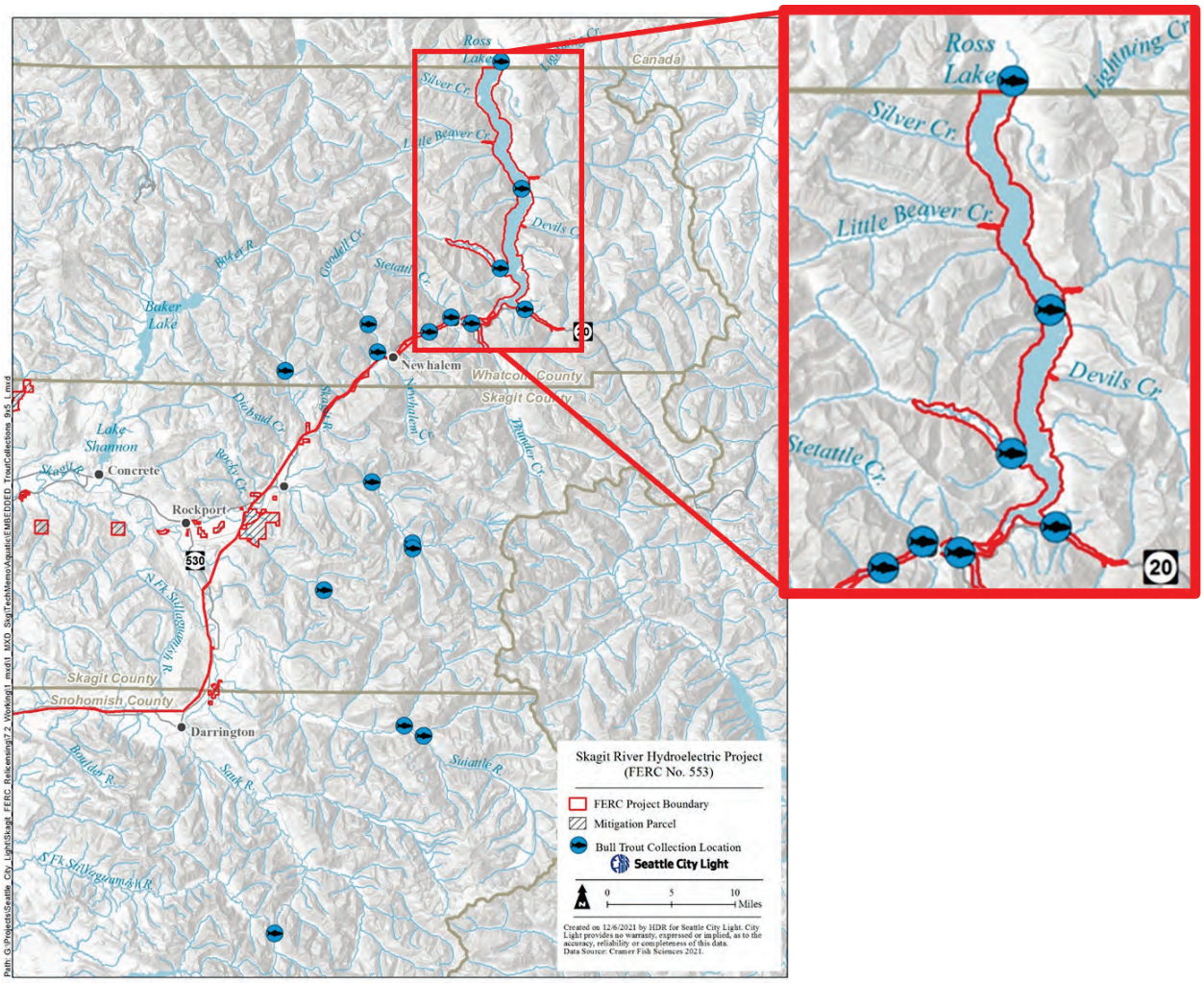


Figure 3.0-2. **Map showing the collection locations of Bull Trout analyzed in Year 1. Year 1 data were provided in an ad hoc fashion and so important information about the specific locations of collections was not available. Some collections were made within the study area (outlined in red), while others were outside of it. For samples collected downstream of Gorge Dam, it is not known whether collections were made upstream or downstream of barriers or within the anadromous zone, because the metadata regarding the sampling locations was not provided. The USFWS samples were assumed to be at the same locations as reported in Smith (2010), but no location data was provided. Samples obtained from within the study area were considered "at-large" from reservoirs, as no location information was available from the WDFW other than the collection code. See Table 5.2-1 for details on anadromous barrier locations identified in Year 2 field surveys.**

4.0 METHODS

4.1 Introduction and Background on Analysis Approach

The methods for this report are divided into two years. In Year 1, the focus was on collecting, standardizing, and verifying existing data to assess the baseline population genetics of salmonids in the Project reservoirs. Year 2 was aimed at filling the data gaps identified in Year 1. For both study years an Expert Panel served an advisory role for the study. Fundamentally, methods in both years were to identify genetic populations and measure their genetic diversity and similarities (i.e., population structure). Although the genetic analysis conducted in both years was similar, Year 2 had more in-depth methods (e.g., Genotyping-in-Thousands by Sequencing [GT-Seq] Single Nucleotide Polymorphism [SNPs] in Year 2 vs. microsatellites in Year 1). Year 1 provided a summary of existing information for discussion purposes, while Year 2 analyzed newly collected samples using updated genotyping techniques.

Defining populations is a complex task due to the intricacies of population structure in nature. This report categorizes individuals into "genetic populations" using mathematical models, but it is important to recognize that delineating populations in the real world is not binary (yes/no), can be complex and nonlinear (e.g., influenced by hybridization), and can range from reproductive panmixia to complete isolation. There are two main methods for defining populations: the ecology-based approach focuses on demographic cohesion (i.e., independence of vital rates) and the evolution-based approach focuses on reproductive cohesion (Waples and Gaggiotti 2006). Thresholds for independence among populations within these two categories are migration rate (m) for the ecological paradigm and migrants per generation (Nm) for the evolutionary paradigm. The relationship between demographic independence and genetic diversity is complex, but there are frameworks to assess the relative independence of populations that consider factors such as the Fixation Index (F_{ST}), random genetic drift, inbreeding, migration rate (m), and adaptation (Lowe and Allendorf 2010). For instance, populations can be designated based on the relatively small effect of genetic drift if F_{ST} between two populations is less than 0.02, which indicates successful migration (gene flow) is sufficient to maintain similar allele frequencies in both populations (Lowe and Allendorf 2010). Another F_{ST}-based approach considers the relationship between F_{ST} and migration rate (m), with populations considered demographically independent if m is less than 10 percent (Lowe and Allendorf 2010). As an alternative to F_{ST}, simulation-based approaches can also be used to evaluate multiple evolutionary forces simultaneously. This report is not meant to review or vet the appropriateness of the different ways to define populations as they apply to native fish in the Project area. Rather, it provides relevant statistics for each designated (modeled) "population" – regardless of the method or threshold that is ultimately used by fishery managers – to support the identification of genetic populations and subsequent discussions about future monitoring needs.

4.1.1 Overview of Analyses

The purpose of each genetic test performed in this study is briefly described below and in Table 4.1-1 to provide context to the reader. The goal of population genetics is to understand the distribution of genetic variation within and among populations and the processes that shape the genetic structure of populations. Population genetics is used to address various questions, such as enumerating populations, determining the origin of fish used for artificial propagation, and evaluating the relationships between populations from different geographic locations or their

viability. To conduct population genetic assessments, it is necessary to first determine the population(s) present in the sample. This is an iterative process that involves combining metadata from collections with evaluations of genetic data models, such as by partitioning genetic data by geographic location or using a statistical model that estimates similarities/dissimilarities among individuals to identify populations. Once populations are identified, genetic diversity within and among populations can be evaluated.

The **Hardy-Weinberg equilibrium (HWE)** is a mathematical model that describes how the frequency of genetic variants (alleles) in a population should remain constant from generation to generation, assuming no evolutionary forces that can alter genetic diversity, such as migration, random genetic drift, selection, or mutation. The goal of testing for HWE is to assess if a population's genetic variation aligns with the predictions of the HWE model. The model predicts how genetic variation in a population should remain constant over generations in the absence of evolutionary forces, such as migration, genetic drift, selection, and mutation. If a population deviates from the predicted relationship between allele frequencies and genotype frequencies, this may indicate that these forces are affecting the population's genetic diversity. Testing for HWE determines if alleles are randomly associated within each genetic marker, which is important for identifying potential issues with non-random mating. It also provides information about the quality of the genetic data, which is crucial for the accurate characterization of Rainbow Trout, Bull Trout, and Dolly Varden populations within the study area. However, it is important to note that deviation from HWE does not necessarily mean that the samples being analyzed are not a population, but rather that the population is not following the predictions of the model. The goal of HWE testing is to identify agreement with the model and to accurately enumerate the populations under study.

Linkage disequilibrium (LD) is the nonrandom association of alleles among genetic markers (loci). Despite the term, genetic markers physically co-located on a chromosome does not necessarily mean alleles at markers would be present together more often than expected, as recombination during cell division breaks association over time. LD is a statistical correlation. Quantifying LD can be of interest as population history is reflected in the distribution of LD genome wide, while localized LD at a genomic region reflects forces acting to alter genetic variation (alleles present). Specific to the Reservoir Fish Genetics Study, observed LD was used to estimate the effective population size. The reproductive process is expected to generate higher LD in a small population relative to a large population. There is a well-defined mathematical relationship between LD and effective populations size.

Importantly, due to the potentially very high number of statistical tests for HWE and LD, significance thresholds are often adjusted using a correction for multiple tests (i.e., the family-wise error rate [FWER]). The Bonferroni correction is a conservative method that is commonly used to control the FWER when conducting multiple statistical tests on a dataset. The Bonferroni correction reduces the likelihood of making a type I error (i.e., rejecting the null hypothesis that genotypes are in HWE and LD when it is true) but at the cost of reducing statistical power. It is a simple and widely recognized method but can result in overly conservative results, especially when the number of tests is large. Additionally, it assumes that the tests are independent, which may not always be the case in practice. Thus, the Bonferroni correction is a useful tool to control the FWER when testing for HWE and LD with multiple loci, but its application should be carefully considered to avoid overly conservative results and maintain statistical power.

Observed heterozygosity (H_O) and **expected heterozygosity (H_S)** are measures of genetic variation in a population. A heterozygote has different alleles at a particular genetic marker (locus). Observed heterozygosity is the proportion of individuals in a population that are heterozygous at a genetic marker. Observed heterozygosity is an empirical measurement of the level of genetic diversity in a population based on the actual alleles present. Expected heterozygosity is the level of heterozygosity that would be predicted under HWE, assuming random mating and no forces were affecting the frequency of alleles in the population. Comparing observed and expected heterozygosity provides insight into the forces acting on genetic variation in the population. If observed heterozygosity is higher than expected heterozygosity, it may indicate the presence of successful reproduction among dissimilar individuals (migration) or a declining effective population size. Conversely, if observed heterozygosity is lower than expected heterozygosity, it may suggest the presence of non-random mating (inbreeding) in the population or selection. Another measure of genetic diversity is allelic richness, which is the average number of alleles present per genetic marker (locus). **Allelic richness (Ar)** is a measure related to a population's viability (potential to adapt and persist).

The effective population size (N_e) is among the most important parameters in conservation because it influences the efficiency of natural selection, gene flow, inbreeding, and loss rate of genetic variation within a population. Ultimately, the N_e is defined by demography, but mathematical relationships between genetic metrics and the N_e can be used to estimate N_e from genetic data. The effective population size can be calculated annually (**N_b; effective number of breeders**) or over a complete generation (N_e) for a single population or for a reproductively connected group of populations (i.e., a metapopulation). Monitoring N_e and N_b can facilitate early detection of population declines or potential inbreeding that relates to a decreasing viability. In the Reservoir Fish Genetics Study, the linkage disequilibrium quantified within juvenile samples from tributaries was used to estimate N_b, the effective population size of parents that gave rise to the particular same-aged cohort in one reproductive cycle. Values of N_e are often interpreted in relation to thresholds of the 50/500 rule. The 50/500 rule is a guideline used in management and conservation to determine the minimum viable population size for a species. The rule states that to ensure the long-term survival of a species, an effective population size of at least 50 individuals is needed to avoid the negative effects of inbreeding and genetic drift (i.e., decreased reproductive success, loss of trait diversity), and a population size of at least 500 individuals is necessary to reduce the risk of extinction from demographic and environmental stochastic events; however, the 50/500 rule is a "rule-of-thumb" and not a strict standard. The actual minimum viable population size may vary depending on many factors (life history, genetics, habitat quality, level of threats).

F-statistics were used to measure population structure. The classical analysis of population structure assumes that the samples being analyzed come from pre-determined subpopulations. Statistical tests using F-statistics (such as F_{IS}, F_{ST}, and F_{IT}) are used to determine if this assumption is accurate. In the Reservoir Fish Genetics Study, **F_{IS}** metrics were used to assess the deviation between observed heterozygosity and expected heterozygosity, indicating if the collections of samples represent populations. It is important to note that F_{IS} measures inbreeding relative to a randomly mating population and should not be confused with inbreeding relative to a shared recent common ancestor (also known as pedigree inbreeding). F_{IS} should therefore not be used to directly assess the risk of inbreeding depression.

In the Reservoir Fish Genetics Study, ***F*ST** metrics were used to quantify the magnitude of genetic differentiation between collections (or populations). The F_{ST} metric was originally described as the ratio of the observed variance of allele frequencies between collections (or populations) to the expected variance of allele frequencies assuming no differentiation from population structure exists (i.e., panmixis). In the Reservoir Fish Genetics Study, the F_{ST} was estimated between populations using an analysis of variance of allele frequencies. Given a defined population structure, a hierarchical analysis of variance partitions total genetic variance into covariance components due to within- and between-population differences. The covariance components are used to estimate F_{ST}. If two collections (or populations) being compared are genetically differentiated ($F_{ST} > 0$), the magnitude of differentiation can be quantified by the F_{ST} statistic, akin to a genetic distance. In the Reservoir Fish Genetics Study, testing of F_{ST} focused on whether the F_{ST} calculated between two collections (or populations) was statistically different from zero.

Interpretation of F_{ST} magnitudes will vary by biological or management objective. Definitions of "populations" typically fall under two paradigms: those reflecting an ecological paradigm (which emphasizes demographic cohesion) and those reflecting an evolutionary paradigm (which emphasizes reproductive cohesion) (Waples and Gaggiotti 2006). The F_{ST} is a commonly used population metric because it has desirable aspects for measuring population independence that are rooted in biological significance rather than arbitrary population designations (e.g., those simply based on rejection of reproductive panmixia; $F_{ST} > 0$). Yet, the ability to reject panmixia (i.e., to say two tributary collections are different) is a question of statistical power and not biological significance. F_{ST} value thresholds that have biological meaning or management relevance can be agreed upon. For example, when F_{ST} between two populations is 0.00 and 0.02, the historical average number of successful migrants exchanged per generation is greater than 10. Within this window, the potential negative effects of diversity loss from random genetic change (genetic drift) are limited because reproductive connection (successful migration) is expected to maintain the presence of alleles ('genetic drift connectivity'; Lowe and Allendorf 2010). This approach is not without limitations, namely, the assumption of migration-drift equilibrium and that populations are known/defined *a priori*. Nevertheless, the metric is routinely calculated and can serve as an index of population structure within an adaptive management framework.

In the Reservoir Fish Genetics Study, given the paucity of information related to char in the Study Area, further context was provided for F_{ST} metrics between char populations by testing for isolation-by-distance and migration-drift equilibrium. **Isolation-by-distance** describes a pattern of population genetic variation derived from spatially limited successful migration (gene flow), where genetic differences between populations increase with increasing geographic distance (i.e., stepping-stone relationship). Isolation-by-distance is consistent with the assumption of **migration-drift equilibrium** and was evaluated using Mantel tests of fluvial distance and genetic distance. The assumption of migration-drift equilibrium was further assessed using a second Mantel test using residuals from the initial fitted correlation line against pairwise riverine distance. At equilibrium, scatter (residuals) should increase with increased geographical separation as random changes (drift) become a more dominant force than migration (gene flow) on genetic diversity at greater geographic distances. Equilibrium conditions between gene flow and drift are common among populations of salmonids, especially among neighboring populations (Hutchison and Templeton 1999). Yet, the expected pattern of genetic diversity in char was unclear because there was significant isolation-by-distance coupled with nonsignificant scatter of residuals (see Section 5, Results). Nonequilibrium/dynamic equilibrium conditions may be expected in recently

deglaciated watersheds due to the complex history of extirpation/recolonization that occurred during the Pleistocene (Taylor et al. 2003). Migration-drift equilibrium is also an implicit assumption of most population genetic mathematical models and provides a useful means to test hypotheses about genetic change (evolution).

The Project team utilized both supervised and unsupervised models to examine genetic structure among different collections. In population genetics, a supervised model involves using prior knowledge of the population structure to make predictions about the population, while an unsupervised model relies solely on the data to make predictions. For the supervised models, the subpopulations were assumed to be the same as the collections, and contemporary watershed boundaries were used to define the subpopulation boundaries (e.g., samples collected from Lightning Creek were treated as a sample from a subpopulation occupying Lightning Creek).

To test for genetic population structure, the samples were assumed to be collected from well-defined subpopulations characterized by the tributaries from which they were collected. The subpopulations were assumed to be part of a metapopulation, whose boundaries were defined by contemporary watershed boundaries. A hierarchical approach was used, in which collections were nested within lakes, and the **Analysis of Molecular Variance (AMOVA)** was used to partition genetic variation within individuals, among individuals within subpopulations, among subpopulations, and among groups (i.e., lakes). For example, Lightning Creek was treated as a subpopulation nested within Ross Lake.

In the Reservoir Fish Genetics Study, population structure was also investigated using a statistical approach that does not require populations be defined *a priori* (i.e., unsupervised data analysis). Unsupervised data analysis is often desirable when the underlying genetic structure among collections or populations is unknown. **Principal components analysis (PCA)** is a multivariate approach to make inferences about genetic population structure. In the Reservoir Fish Genetics Study, principal components analysis focuses on shared genetic variance, summarizing multiple variables (alleles) with the minimum number of components such that each component explains the most genetic variance (i.e., reduces the dimensionality of data). Principal components analysis creates new, uncorrelated variables that successively maximize explained variance. In the event of finding such new variables, the principal components (PCs), reduce to solving an eigenvalue/eigenvector problem, and the new variables are defined by the dataset at hand, and not populations defined by the user *a priori*. Hence PCA is an adaptive and unsupervised data analysis technique. A key output of PCA is the inertia, which refers to the sum of squared distances of all the points in a dataset to the origin. It measures the amount of variation in the data that is captured by the PCs. The inertia of a PCA provides information about the total variation in the data and how much of that variation is explained by each PC. Interpreting the inertia can help determine the appropriate number of PCs to retain for further analysis and gain insights into the structure of the data. Thus, fish with relatively high/low inertia values can be interpreted as being genetically distinct.

Yet, because PCA describes total genetic variance (diversity) in the dataset, differences between groups within the dataset may be obscured. For example, the genetic diversity within the identified groups may not confer to population model expectations (e.g., HWE). In contrast, **discriminant analysis of principal components (DAPC)** is designed to identify and describe clusters of genetically related individuals, enhancing resolution among groups defined by genetic variance.

DAPC uses information from all genetic markers to create new axes and then projects the data in a way that maximizes separation of genetic groupings. DAPC uses sequential k-means and Bayesian Information Criterion (BIC) model selection to infer genetic clusters (i.e., the prior). K-means clustering is a method of vector quantization, originally from signal processing, that aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype (i.e., prior) of the cluster. Membership probabilities are then estimated using retained discriminant functions (i.e., posterior probabilities of assignment). DAPC is often preferred over a STRUCTURE-like approach because it does not assume panmixia within genetic clusters and can accommodate more complex population relationships such as hierarchical or steppingstone (Jombart et al. 2011). In the Reservoir Fish Genetics Study, populations in the Study Area were identified using an iterative process, with groups identified using DAPC analyzed using population models, which in turn refined subsequent DAPC and population analysis.

The study team chose DAPC as the method for analyzing population structure because it has relatively few restrictions on model assumptions relative to "STRUCTURE-like" algorithms. Although some members of the EP preferred using STRUCTURE, the team chose DAPC because it is less restrictive. **STRUCTURE** is a Bayesian clustering algorithm that uses pre-defined models of HWE and Linkage Disequilibrium. To gain a better understanding of population clustering, the genetic dataset was analyzed with STRUCTURE and the results were compared qualitatively to those from DAPC. However, since DAPC and STRUCTURE are different in their analytical approach, no direct statistical comparisons were performed.

The identification of parent-offspring pairs (parentage) or full-sibling pairs (relatedness) can provide useful biological monitoring information. Parentage can determine whether fish move between reservoirs and their subsequent survival, gauge reproductive success within reservoirs, or estimate population abundance. Populations with a small effective size are expected to be more related on average than populations with large effective size. In the absence of effective size information, distribution of relatedness among individuals can be used to inform biological risk assessment (i.e., potential rate of genetic diversity loss). In the Reservoir Fish Genetics Study, base condition **relatedness** among individuals was quantified using the statistically unbiased Queller and Goodnight (1989) Rxy estimator. Further, should parentage methods be included in a future monitoring plan, the statistical reliability should be determined. In Year 1 the statistical power (**false detection rate**) of existing data (microsatellites) to identify relationships among individual Bull Trout and Rainbow Trout (e.g., parent-offspring pairs, full-sibling-unrelated pairs) was also determined. Note that the statistical power of existing data was low and determined to be insufficient for monitoring applications. This Year 1 determination provided justification to update genotyping methods in Year 2 to single nucleotide polymorphism genetic markers.

Table 4.1-1. Glossary of terms, statistical test, test metric, purpose of test, significant value threshold, species tested.

Statistical Test	Metric	Purpose	Threshold	Species
Hardy-Weinberg equilibrium (HWE)	F_{IS}	Data QA/QC and enumeration of populations in Study Area. Determine if a population's genetic variation agrees with the predictions of the Hardy-Weinberg equilibrium model.	Probability of samples conditional on observed allelic counts (per locus exact test); null rejection often 0.05 α corrected for multiple tests	Rainbow Trout Bull Trout Dolly Varden
Linkage disequilibrium (LD)	R^2; D	Enumeration of populations in Study Area. Determine if a population's genetic variation exhibited statistical correlations between genetic markers.	Probability of samples conditional on observed genotype counts (log likelihood ratio statistic G-test); null rejection often 0.05 α corrected for multiple tests	Rainbow Trout Bull Trout Dolly Varden
Observed heterozygosity	H_O	To measure the observed proportion of heterozygotes in a collection of individuals. Can be compared to the expected heterozygosity in a test of HWE.	Variable; often 0.05 or 95% confidence interval (CI)	Rainbow Trout Bull Trout Dolly Varden
Expected heterozygosity	H_S	To measure the expected proportion of heterozygotes in a collection of individuals from a population based on Hardy-Weinberg Equilibrium.	Variable; Often 0.05 or 95% CI	Rainbow Trout Bull Trout Dolly Varden
Allelic richness	Ar	To measure the number of different alleles in a population corrected for sample size using rarefaction.	NA	Bull Trout Rainbow Trout
Effective population size	N_e	The size of an ideal population (one in HWE) that would lose diversity at the same rate as the population being considered. It is a theoretical number and it is not equivalent to the number of adults that successfully reproduce.	Uses mathematical relationship between Linkage Disequilibrium and N_e $N_e > 50$ to avoid negative effects of inbreeding $N_e > 500$ to reduce the risk of extinction from stochastic events	Rainbow Trout Bull Trout Dolly Varden
Effective number of breeders	N_b	To measure single cohort N_b which provides a metric of population-specific individual reproductive contribution.	Uses mathematical relationship between Linkage Disequilibrium and N_e	Rainbow Trout Bull Trout Dolly Varden
F-statistics	F_{IS}	Used to evaluation Hardy-Weinberg equilibrium model. Estimates the departure from panmixia at the level of the subpopulations.	Expressed using ANOVA components of variances $\hat{\sigma}_G^2$, $\hat{\sigma}_I^2$	Rainbow Trout Bull Trout Dolly Varden

Statistical Test	Metric	Purpose	Threshold	Species
F-statistics	F_{ST}	Enumerate populations in Study Area. Quantified the magnitude of genetic differentiation (in allele frequencies) between collections (or populations);	Expressed using ANOVA components of variances $\hat{\sigma}_G^2$, $\hat{\sigma}_I^2$, $\hat{\sigma}_P^2$ Define population thresholds: F_{ST}<0.02 "drift connectivity"; F_{ST} < 0.20 "inbreeding connectivity"; F_{ST}<0.35 "adaptive connectivity"	Rainbow Trout Bull Trout Dolly Varden
Isolation-by-distance	R^2	Determine if there was a statistical correlation between genetic and geographic distance; evaluate migration-drift equilibrium	P-value; R^2; Regression; Mantel	Dolly Varden
Migration-drift equilibrium	Null model expectation	Serves as the null hypothesis in most population genetics studies. Used to provide insight into the forces that drive evolution and shape the distribution of genetic variation in populations.	Regression Mantel	Dolly Varden
Principal components analysis (PCA)	F_{ST} genetic distances	To identify the underlying variance structure of complex data by reducing the dimensionality of the data and identifying the most important underlying patterns	NA	Rainbow Trout Bull Trout Dolly Varden
Discriminant analysis of principal components (DAPC)	k	To test the significance of the principal components in classifying individuals into predefined groups. Does not assume HWE.	Bayesian Information Criterion (BIC) model selection; Posterior probabilities of membership probabilities estimated using retained discriminant functions	Rainbow Trout Bull Trout Dolly Varden
STRUCTURE	K	To infer the genetic structure of populations and the relationships between individuals within and between populations. Assumes HWE.	q-value, which is the estimated proportion of an individual's genome derived from a specific taxon.	Rainbow Trout
Relatedness	R_{xy}	Evaluate parentage as a genetic monitoring option. Metric quantified the relatedness between two individuals.	$R_{xy \sim}$ 0.00 unrelated R_{xy} > 0.25 half siblings (share one parent) R_{xy} > 0.50 full siblings (share two parents)	Rainbow Trout Bull Trout

Statistical Test	Metric	Purpose	Threshold	Species
Parentage statistical power	false detection rate	Evaluate parentage as a genetic monitoring option. Statistical power to correctly infer relatives from unrelated individuals.	False positive rates when the true relationship is unrelated. False negative rates when the true relationship is related. General recommendation for being confident about not erroneously identifying unrelated individuals as related pairs is to require that the FPR be about 10 to 100 times smaller than the reciprocal of the number of comparisons.	Rainbow Trout Bull Trout Dolly Varden

4.2 Data Requests and Sample Collections

4.2.1 Year 1 Data Requests and Sample Collections

During year 1, City Light identified the availability of existing genetic samples from past studies (e.g., unanalyzed samples from past studies, archived samples from fieldwork in the study area, samples used in previous analyses for which a partial sample may still be available for additional analyses, etc.). City Light also coordinated potential opportunistic sampling conducted by other relicensing studies and current Project license-related field activities. These additional sampling opportunities included: the U.S. Geological Survey (USGS) Food Web Study (Beauchamp, in development),[2] the Acoustic Telemetry Monitoring Program, and the FA-03 Reservoir Fish Stranding and Trapping Risk Assessment (City Light 2023a).

Data requests were made to state and federal researchers to obtain the previously identified and pertinent microsatellite genotypes listed in Table 2.5-1 of the RSP. The datasets contained existing genotypes that may be useful for informing Year 1 objectives. On June 6, 2021, Cramer Fish Sciences (CFS) emailed Todd Seamons, Director of the WDFW genetics laboratory, and Matt Smith, fish geneticist at the USFWS Abernathy Fish Technology Center, requesting data and metadata used in Pflug et al. (2013) and in Smith (2010).

On June 13, 2021, Matt Smith with USFWS provided (via email) a tab-delimited .txt file containing 563 *Salvelinus* genotypes at 16 microsatellite loci used in Smith (2010): *Omm1128*, *Omm1130* (Rexroad et al. 2001); *Sco102*, *Sco105*, *Sco106*, *Sco107*, *Sco109* (WDFW unpublished); *Sco200*, *Sco202*, *Sco212*, *Sco215*, *Sco216*, *Sco218*, *Sco220* (Dehaan and Ardren 2005); *Sfo18* (Angers and Bernachez 1996); and *Smm22* (Crane et al. 2004). The dataset sent by Matt Smith included the following metadata: Individual Name, Synonym 1, Region (1), Watershed (2), Tributary (3), Capture Location (4), Age, Brood Year, Collected By, Collection Year, Comment, Date Collected, Fork Length (millimeter [mm]), Hatchery/Wild, hatchery origin (HOR)/natural origin (NOR) Assignment, Latitude, Life History, Stage, Longitude, Phenotypic Sex, PIT Tag, Population ID, Preservation Method, Project Number, Received From, Resident / Anadromous, Run Type, Spawn Date, Spawn Year, Spawned With, Species, Synonym 2, Synonym 3, Tissue Type, Total Length (mm), Used for Broodstock, and Weight (g). Only some of these metadata were relevant to this study report or contained entries.

On July 28, 2021, WDFW provided (via email) an Excel spreadsheet containing 335 *Salvelinus* and 2,967 *Oncorhynchus* genotypes. The *Salvelinus* were comprised of six collections from reservoirs in the study area and two collections from outside the study area, with genotypes generated using the same microsatellite loci used by Smith (2010). The *Oncorhynchus* genotypes were the 15 microsatellites analyzed in Pflug et al. (2013): *One-102, Ogo-4*, (Olsen et al. 1998); *Ots-100* (Nelson et al. 1998); *Oki-10*, *Oki-23* (Smith et al. 1998); *Omy-7* (K. Gharbi, unpublished, as referenced in Pflug et al. 2013*); Omy-1001*, *Omy-1011* (Spies et al. 2005); *Ots-3M*, *Ots-4* (Banks et al. 1999); *One-14* (Scribner et al. 1996); *Ssa-407, Ssa-408* (Cairney et al. 2000); *Ssa-298* (McConnell et al. 1995); and *Oke-4* (Buchholz et al. 2001). The dataset included the following metadata: Sample ID, WDFW Collection Code, Count, and Percent Missing Data. Various other

[2] The Food Web Study is an ongoing voluntary study (outside the FERC-approved study plan) developed in consultation with the Flow/Non-Flow Committee and initiated prior to the Project relicensing proceedings. The findings of the Food Web Study will be published in a series of USGS manuscripts in 2023.

metadata were available directly from the Pflug et al. (2013) and Smith (2010) reports. No geospatial data was available.

4.2.2 Year 2 Data Requests and Sample Collections

The results from Year 1 were summarized and shared with both the LPs and the Expert Panel. This collaboration allowed for the informed planning of Year 2, which was reflected in the table note for Table 4.2-1. The Expert Panel was also consulted at each of the three milestones proposed (i.e., the three Expert Panel meetings), in accordance with RSP 2.5.1 and 2.8. This was further emphasized by the June 9 Notice to seek the Expert Panel's advice on how genetics information and other monitoring methods could be used to evaluate reservoir fish abundance, habitat use, and migration timing in the future. The Expert Panel's consultations were a crucial step in the methodology and were revisited throughout to provide a comprehensive understanding of the process.

Following presentation of and discussion on preexisting genetic data with the Expert Panel and LPs (Year 1 Technical Memo; Attachment A hereto), Year 2 data requests and sample collection activities were refined. The February 9, 2022, Expert Panel independent working session and the March 31, 2022, meeting with the Expert Panel and LPs discussing the data, approach, questions, and objectives of the Reservoir Fish Genetics Study were informative for Year 2 planning (Details of Expert Panel deliberations provide in Section 5.0, Results section).

In Year 2 (2022), City Light focused efforts on obtaining the samples and genotypes needed to characterize genetic population structure and relationships and to begin estimating N_e (to evaluate population viability). The key consideration to sampling for population structure and effective population size is to acquire random samples from the cohorts that comprise the generation(s) of interest within representative subpopulations of the metapopulation (Ryman et al. 2019). Tissues and SNP genotypes were requested from USGS and WDFW to augment the new tissues collected from the study area during summer 2022. USGS provided 663 *O. mykiss* tissues sampled between 2018 and 2021 from across ten tributaries. WDFW provided genotypes for N=30 *O. mykiss* from Gorge Lake containing data comprised of the same GTSeq panel genotypes generated by CFS for this study and N=428 *O. mykiss* from nearby below-Project (i.e., below Gorge Dam) locations containing a subset of GTSeq panel genotype data. The WDFW genotype data was primarily used for comparisons of *O. mykiss* within and outside of the study area. WDFW also provided 180 tissues from Dolly Varden collected within the study area between 2019 and 2020 (Seamons 2020). Seattle City Light provided an additional 32 Bull Trout tissues collected from the Project reservoirs between 2020 and 2022 (Fisher 2020).

Table 4.2-1 lists the sampling locations targeted for the 2022 field season, which were believed to be representative of the subpopulations contributing to the productivity and genetic diversity of native trout/char in the study area based on eDNA surveys and previously identified char spawning locations (Ostberg 2022). It is important to note that the sampling locations in Year 2 (summer/fall 2022) were the same for *Oncorhynchus* and *Salvelinus*. However, not all sites contained both genera during electrofishing surveys. For instance, Little Beaver Creek contained *O. mykiss* but no *Salvelinus*, which is why certain sites are listed for one taxon and not for the other.

Despite the challenges during the field season, such as an active wildfire season and access issues due to safety concerns, a robust genetic baseline was still obtained. Fourteen tributaries were

sampled from Ross Lake, both tributaries proposed for Diablo Lake, and 2 of the 3 tributaries for Gorge Lake were sampled. This resulted in a robust genetic baseline for the Reservoir Fish Genetics Study (Table 4.2-1, Figure 4.2-1), which also displays the natural barriers to fish migration measured and determined during the FA-07 Reservoir Tributary Habitat Assessment Study.

Table 4.2-1. Summary of locations[1] targeted for sampling *Oncorhynchus* and *Salvelinus* during the Year 2 field season in 2022.

No.	Skagit River Drainage	River/Stream Name	Location Description	Sampled in 2022
1	Ross Lake	Hozomeen Creek	Mainstem	Yes
2		Freezeout Creek	Mainstem	No
3		Lightning Creek	Mainstem	Yes
4		Three Fools Creek	Mainstem	Yes
5		Cinnamon Creek	Mainstem	No
6		Castle Fork	Mainstem	No
7		Devils Creek	Mainstem and tributaries	No
8		North Fork Devils Creek	Mainstem	No
9		Roland Creek	Mainstem	Yes
10		Ruby Creek	Mainstem	Yes
11		Canyon Creek	Mainstem up to cascade barrier	Yes
12		North Fork Canyon Creek	Mainstem	No
13		Granite Creek	Mainstem up to cascade barrier	Yes
14		Panther Creek	Mainstem up to cascade barrier	Yes
15		Pierce Creek	Mainstem	Yes
16		Big Beaver Creek	Mainstem and tributaries including Beaver Ponds	Yes
17		McMillan Creek	Mainstem	Yes
18		Luna Creek	Mainstem	Yes
19		Little Beaver Creek	Mainstem above and below barriers	Yes
20		Silver Creek	Mainstem	Yes
21	Diablo Lake	Thunder Creek	Mainstem	Yes
22		Colonial Creek	Mainstem	Yes
23	Gorge Lake	Stetattle Creek	Mainstem above and below barrier	Yes
24		Pyramid Creek	Mainstem above barrier	Yes
25		Gorge Creek	Mainstem	No

1 The proposed sampling plan for 2022 field study, including sampling locations, was reviewed and approved by the Expert Panel and interested LPs. Locations include those sampled by CFS and USGS.

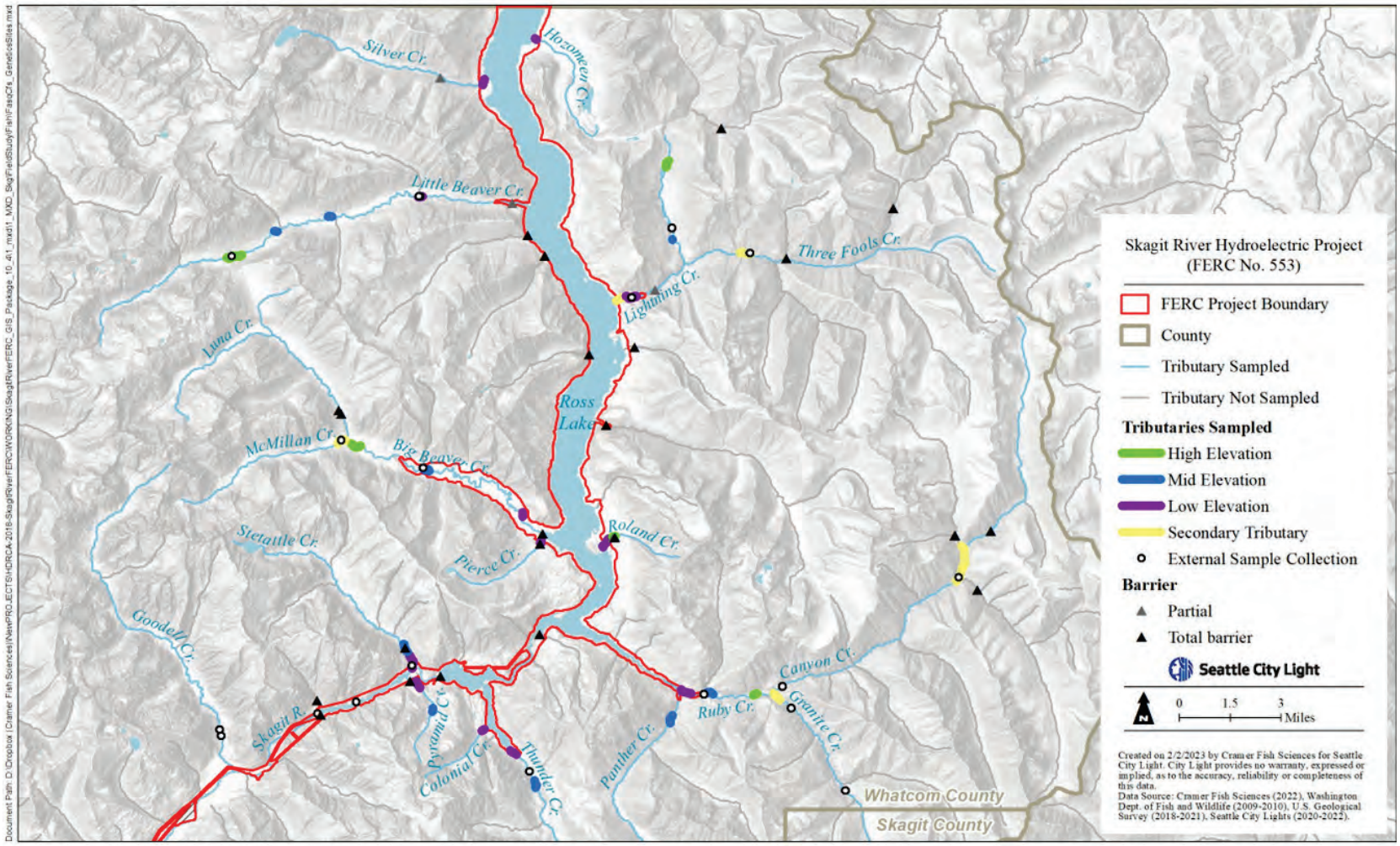


Figure 4.2-1. **Year 2 *Salvelinus* and *Oncorhynchus* genetics sampling locations. Primary tributaries were identified by direct connectivity to the reservoirs and divided into three sampling strata, a low, mid, and high elevation, to avoid overrepresentation of the family groups. Secondary tributaries are identified by connectivity to a primary tributary, do not flow directly into the reservoirs, and were only sampled at a single location. Locations such as Silver mid and high, Thunder high, Colonial mid and high, were not sampled due to challenges during the field season. Other locations, such as Luna Creek and Hozomeen Creek, were sampled but no fish were collected.**

To describe genetic diversity and structure, the study team sought to collect representative genetic samples from fish occupying the major spawning and nursery grounds of rivers and streams in the Project area. The team chose to target trout fry (young of year [YoY]) because this is the early life history stage directly related to reproduction and early development habitats (Garant et al. 2000). Sampling fry within spawning/nursery habitats provides insight into the distribution of adult trout and char returning to their natal spawning sites and is likely to reflect the true population structure (Garant et al. 2000). However, sampling emergent fry may increase the likelihood of finding statistically significant genetic differentiation among sites due to overrepresentation of family groups (i.e., the Allendorf-Phelps effect). Sampling was therefore conducted using an adjusted version of the recommendations of Whiteley et al. (2012) to avoid overrepresentation of family groups. Specifically, sampling targeted three distinct locations (i.e., high, mid, and low elevation) in each tributary and collections were made via backpack electrofishing, following guidelines in Fisheries Techniques (Reynolds and Kolz 2012), and according to Institutional Animal Care and Use Committee permitting conditions.

Tissues and scales (to age fish for N_b analysis) were collected from up to 50 individuals per tributary. A sample size of 50 is expected to provide enough power to detect an effective size between 50 and 500 (i.e., in reference to the "50/500 rule"[3] [Franklin 1980; Waples 2006; and Whiteley et al. 2012]). Yet, due to the possibility of encountering fewer than 50 individuals, sampling was scaled by a predetermined survey effort (e.g., initial presence/absence survey, collection of 50 individuals, or up to 90 minutes of electrofishing). Except for fish that appeared to be Eastern Brook Trout (*Salvelinus fontinalis*), no effort was made in the field to collect one taxon over another within each genus (i.e., *Salvelinus* and *Oncorhynchus*) due to phenotypic similarities between species and/or hybrids. This approach of not making efforts to sample any taxon over another is intended to support unbiased inferences about patterns of hybridization across the study area. There was an exception for Eastern Brook Trout because they can be present in high densities, and so the maximum sample size of 50 within *Salvelinus* (N=50) could plausibly be reached early in the sampling effort, thus reducing chances that a representative sample of native *Salvelinus* (i.e., Bull Trout or Dolly Varden) genes is achieved. Therefore, only the first 30 apparent Eastern Brook Trout were sampled (15 within a reach), allowing for an additional 20 samples for all other *Salvelinus*.

4.3 Year 1 Genetic Analysis

4.3.1 Rainbow Trout

Summary characterizations of existing Rainbow Trout genetic data in Year 1 used the program FSTAT Version 2.9.3.1 (Goudet 1995) to estimate and test metrics of genetic diversity unless otherwise stated. Expected H_s and Ar were estimated to describe genetic diversity across loci and collections. Randomization tests were performed to test the assumption of HWE at each locus within collections. Observed heterozygosity (H_O) and H_S multilocus heterozygosity within populations were compared using Wright's (1951) inbreeding coefficient (F_{IS}) to measure the magnitude of departures from HWE. To assess the assumption of random association of alleles among loci, log-likelihood ratio tests using 1,000 permutations were implemented to test for

[3] The "50/500" rule is a rule of thumb in conservation biology that advises a minimum effective population size of 50 is necessary to reduce the negative effects of inbreeding in the short term and a minimum of 500 individuals was needed to reduce genetic drift in the long term.

pairwise LD within all collections. The Weir and Cockerham (1984) version of genetic divergence (F_{ST}) was estimated to measure genetic differentiation between all pairs of collections. A PCA of individual-based genetic distances was implemented using the R package {adegenet} (Jombart 2011) to summarize the genetic diversity among the sampled individuals.

The computer program POWSIM Version 4.1 (Ryman and Palm 2006) was used to estimate statistical power to detect deviation from genetic homogeneity. POWSIM is a simulation-based computer program that estimates statistical power of rejecting the null hypothesis of genetic homogeneity for different combinations of sample sizes, number of loci, number of alleles, and allele frequencies for a hypothetical degree of true differentiation (quantified as F_{ST}). POWSIM can only accommodate 30 collections of individuals, so the first 30 collections were used to estimate power to detect different magnitudes of genetic differentiation (F_{ST}=0.001; F_{ST}=0.01) by assuming allele frequencies estimated from the loci described in this report. The statistical power to observe relatives was determined using close kin mark recapture {CKMRSim} (Anderson 2019). All tests of significance were assessed at the $\alpha = 0.05$ level and applied Bonferroni corrections when conducting multiple tests.

4.3.2 Bull Trout

Exploratory analyses were conducted on Bull Trout like those described for Rainbow Trout. Partitioning of genetic variation was explored using visualization of individual-based data and genetic PCA (e.g., Jombart et al. 2010). The statistical power to observe relatives was determined using {CKMRSim} (Anderson 2019). Tests of genetic equilibrium were performed on collections. Following exploration of genetic data present in collections, summary statistics were calculated. Gene diversity (the expected frequency of heterozygotes within a population assuming HWE) was estimated following the sampling bias correction method described by Nei (1987). The observed heterozygosity (average frequency) was also estimated. A common implementation of the HWE test was used following the Guo and Thompson (1992) Markov-chain random walk extension of Fisher's (2-allele) classical exact test. Departures from HWE were also quantified using the inbreeding coefficient (F_{IS}) statistic observed from AMOVA (Excoffier et al. 1992; Yang 1998), which is equivalent to Weir and Cockerham (1984) small f statistics. Collections were analyzed for evidence of LD (i.e., non-independence of alleles at different loci). LD between a pair of loci was tested using a likelihood-ratio test, whose empirical distribution is obtained by a permutation procedure (e.g., Excoffier and Slatkin 1998). Lastly, allelic distributions across collections were evaluated using contingency table analysis of observed allelic distributions described by Raymond and Rousset (1995).

The AMOVA framework estimates hierarchical f-statistics for any number of desired levels (e.g., within individuals, within populations, among populations). This framework allows for population differentiation (allele frequency variance) to be quantified—the degree that individuals within a population (collection) are more like each other than are individuals from different populations (collection). There are many formulations of the population differentiation variance component measure, although a common implementation is a form of the fixation index (e.g., F_{ST}). In year 1,

estimates of F_{ST}[4] were estimated pairwise following Weir and Goudet (2017) and used as a measure genetic divergence, with statistical significance of F_{ST} metrics from zero calculated following likelihood-ratio tests (Goudet et al. 1996).

4.3.3 Lineage Relationships

For Bull Trout and Rainbow Trout populations, the statistically unbiased Queller and Goodnight (1989) Rxy estimator was used. The power (false detection rate) of genetic markers used to identify relationships among individual Bull Trout and Rainbow Trout (e.g., parent-offspring pairs, full-sibling-unrelated pairs) was also estimated.

4.4 Year 2 Genetic Analysis

4.4.1 Year 2 Laboratory methods

Genomic DNA was extracted from fin tissue using Qiagen DNeasy 96 Kits on the Qiagen Qiacube following manufacturer's recommendations and eluted in 200 µL polymerase chain reaction (PCR)-grade water. Extractions were then concentrated via desiccation and re-elution into 15 µL buffer AE. One well of each 96-well plate remained empty to be processed as a "no-template" control. All *Oncorhynchus* samples were genotyped at a panel of 354 SNPs developed by the Columbia River Inter-Tribal Fish Commission (CRITFC) (Hess et al. 2018) using the GTseq method (Campbell et al. 2014). The panel consists of 242 presumably neutral loci, 112 loci linked to putative adaptive genetic variation, three species-diagnostic loci to differentiate *O. mykiss*, *O. clarkii*, and their hybrids, and one sex identification locus. All *Salvelinus* samples were genotyped at a panel of 264 SNPs developed by Bohling et al. (2021) using the GTseq method. The panel consists of 235 presumably neutral loci, 20 species-diagnostic loci to differentiate *S. confluentus*, *S. fontinalis* and their hybrids (these 20 markers begin with the prefix "*sf*" in Table 4.4-1), eight species-diagnostic loci to differentiate *Salvelinus* taxa (*S. alpinus*, *S. fontinalis*, *S. confluentus*, *S. malma*, *S. namaycush*, and *S. leucomaenis*), and one sex identification locus. For *Oncorhynchus* and *Salvelinus* genera, library preparation methods described in Campbell et al. (2014) were followed. Once size-selected, libraries were Qubit-quantified using the Qubit 1x dsDNA HS Assay Kit, normalized to 4 nM and pooled at equal volumes for sequencing on the Illumina MiSeq (MiSeq® Reagent Kit v3 150 cycle). No more than three libraries were pooled and sequenced at a time. Individuals were genotyped with a custom perl script (GTseq_Genotyper.pl; Campbell et al. 2014), and samples were removed if missing data at more than 65 percent of loci.

The Bull Trout GTseq SNP markers were ascertained specifically to analyze genetic variation within Bull Trout, not within Dolly Varden. Thus, there is a possibility for the GTseq SNP markers to provide biased inferences about genetic variation within Dolly Varden (i.e., ascertainment bias). Following identification of Dolly Varden using the taxon-diagnostic markers analyzed in the GTseq panel (see Section 4.3.2), samples were genotyped using an eight-locus microsatellites panel described by Melnik et al. 2020. PCR was performed with 10 µL reaction volumes and Qiagen PCR components. All thermal cycling was conducted using a Bio-Rad C1000 Touch

[4] F_{ST} is a measure in population genetics that quantifies the degree of differentiation between populations by calculating the proportion of total genetic variance due to differences between populations. F_{ST} values range from 0 to 1, with higher values indicating greater differentiation and lower values indicating greater genetic similarity within populations. F_{ST} is not the only measure of differentiation, but it is easily interpreted and useful in identifying populations that are genetically distinct and can be applied in various areas of conservation biology.

Thermal Cyclers. All PCR products were visualized by electrophoresis on an ABI 3730 automated capillary sequencer (Applied Biosystems) contracted from UC Davis Veterinary Genetics Laboratory. Fragment size analysis was completed using Geneious bioinformatics software (Biomatters, Inc., San Diego, California) consistent with the knowledge base available for these genetic markers. Samples were analyzed independently by two people to reduce process errors, with discrepancies in genotype results resolved using consensus. Individuals were retained for analysis if their multi-locus genotypes consisted of at least 6 (or eight) loci.

4.4.2 Year 2 Hybridization

For identification of *O. mykiss* within collections of *Oncorhynchus* from the study area, three diagnostic loci present in the GTSeq panel were used (Oclgshpx357, Omymyclarp404111, and OmyOmyclmk43896). A hybridization index was created that counted the number of Cutthroat Trout *(O. clarkii*) diagnostic alleles present within individual genotypes, which ranged from 0 (nonhybrid) up to 6 (contained only Cutthroat Trout alleles). Any individual with one or more Cutthroat Trout allele was identified and omitted from analysis. As the primary task was to characterize non-migratory *O. mykiss* populations within the study area, investigation of Cutthroat Trout hybridization was 1) beyond the scope of this study and 2) is being undertaken by the NMFS, although the distribution of hybrids observed within the study area is reported.

For the identification of species and hybrids within the *Salvelinus* field collections, type specimens were obtained from each taxon (Bull Trout, Dolly Varden, and Brook Trout) and genotyped using 22 of taxon-diagnostic SNPs described above (Bohling et al. 2021). Twenty-one of the SNPs were diagnostic for Brook Trout and just one SNP was diagnostic for Bull Trout. Dolly Varden was deduced logically by comparison to the 13 Dolly Varden type specimens. The type specimens were provided by WDFW and included 13 Dolly Varden from the Project area, one Brook Trout of unknown origin, and one Bull Trout from the Baker River basin. The diagnostic genotype profiles were used as follows: samples that were heterozygous at any of the 22 diagnostic markers were considered hybrids and removed from the analysis. The remaining putatively nonhybrid samples were then compared to the Bull Trout diagnostic SNP *Salv_SNP_013*. If homozygous TT, it was determined to be a Bull Trout. Genotypes were then compared to *Salv_SNP_008*, and if homozygous TT, the sample was considered a Brook Trout. All other samples that did not match these criteria and had the same genotype profile as the 13 Dolly Varden type specimens, were a Dolly Varden. Notably, all Dolly Varden failed to amplify at *sf000508_CT_HYB*. *Salvelinus* categorized in this way are depicted in the PCA scatterplot shown in Figure 5.3-18. Three genetic clusters reflecting Bull Trout, Dolly Varden, and hybrids are visually apparent with hybrids falling in between the three clusters.

Table 4.4-1. The 22 diagnostic SNP markers for taxa within *Salvelinus* (Bohling et al. 2021). These markers were used alongside field samples to distinguish among Bull Trout, Brook Trout, and Dolly Varden collections from the Project area.

Assay Name	Bull Trout	Brook Trout	Dolly Varden
Salv_SNP_008	A:A	T:T	A:A
Salv_SNP_013	T:T	G:G	G:G
sf000151_AT_HYB	T:T	C:C	C:C
sf000157_01AT_HYB	A:A	C:C	C:C
sf000382_AG_HYB	T:T	A:A	T:T
sf000508_CT_HYB	C:C	G:G	0
sf000559_AG_HYB	A:A	G:G	G:G
sf000754_AC_HYB	C:C	G:G	C:C
sf001164_02GT_HYB	T:T	A:A	A:A
sf002131_AG_HYB	A:A	G:G	A:A
sf002792_01AG_HYB	G:G	C:C	G:G
sf003611_AC_HYB	A:A	G:G	A:A
sf004651_AG_HYB	G:G	A:A	0
sf005440_AG_HYB	G:G	C:C	0
Sfo_12199_79192_HYB	T:T	A:A	A:A
Sfo_2714_25693_HYB	A:A	T:T	A:A*
Sfo_3881_34908_HYB	T:T	A:A	T:T
Sfo_4699_39079_HYB	T:T	A:A	A:A
Sfo_4701_39083_HYB	A:A	C:C	C:C
Sfo_5504_43035_HYB	C:C	T:T	C:C
Sfo_579_12874_HYB	T:T	C:C	T:T
Sfo_9883_66689_HYB	T:T	C:C	T:T

*Only amplified in Dolly Varden about 50 percent of the time

4.4.3 Year 2 Genetic Diversity within Collections

Assumptions that genotypes conformed to HWE proportions and gametic (linkage) equilibrium were tested using exact tests. For the Markov chain parameters, 100,000 dememorizations, 100 batches, and 1000 iterations per batch. Expected heterozygosity and measured deviation from the observed heterozygosity was estimated (H_o) using Wright's (1951) inbreeding coefficient F_{IS}. For Bull Trout, the AMOVA method of partitioning genetic variation by placing samples into collections that reflect contemporary reservoir structure was also utilized. Statistical significance was assessed at the α=0.05 level and was corrected for multiple tests using the sequential Bonferroni method (approximately 33 tests within each collection for Bull Trout and eight within collections for Dolly Varden) (Rice 1989).

4.4.4 Year 2 Genetic Divergence among Collections

4.4.4.1 Pairwise F_{ST} between collections

Pairwise F_{ST} was estimated between each collection using an AMOVA framework (Weir and Cockerham 1984). Isolation-by-distance was evaluated using Mantel tests of pairwise fluvial distance and genetic distance. Genetic distance was calculated as linear F_{ST} in GenAlex. The assumption of migration-drift equilibrium was further assessed using a second Mantel test using residuals from the initial fitted line against pairwise fluvial distance. At equilibrium, scatter (residuals) should increase with increased geographical separation as drift, as opposed to gene flow, becomes the dominant force at greater distances. Equilibrium conditions between gene flow and drift are common among populations of salmonids and reflect the balancing of loss of alleles due to drift against their replacement via gene flow, especially among neighboring populations (Hutchison and Templeton 1999).

4.4.4.2 Discriminant Analysis of Principal Components

Discriminant Analysis of Principal Components was used to provide an unsupervised analysis of the genetic structure of Bull Trout, Dolly Varden, and Rainbow Trout in the study area. The R package {adegenet} was used to implement the DAPC, beginning with the function *find.clusters()* to identify clusters (i.e., the clusters represent prior population groupings). Where *k* is the number of prior clusters, from *k*=2 up to 20 were explored, including all PCs. The function "*find.clusters*" produces a scatterplot showing the relationship between the BIC (a measure of model fit) and the values of *k*. The relationship between BIC and the best *k* (i.e., the *k* that explains the highest level of structure) depends on the true population genetic structure, which is assumed to be unknown. Thus, the rule of thumb advises increasing *k* until BIC no longer leads to an appreciable improvement of fit (i.e., to a decrease of BIC). In the simplest models (island models), BIC decreases until it reaches the optimal *k*, and then increases. In these cases, the rule amounts to choosing the lowest *k*. under the steppingstone model, however, BIC often continues to decrease after the optimal *k*, but is much less steep. The a-score was estimated to avoid overfitting the models, which is simply the difference between the proportion of successful reassignment of the analysis (observed discrimination) and values obtained using random groups (random discrimination) (Jombart 2011). It can be seen as the proportion of successful reassignment corrected for the number of retained PCs.

4.4.4.3 Year 2 STRUCTURE Analysis

STRUCTURE was used to examine the population structure within the Study Area *O. mykiss* dataset to qualitatively compare with DAPC results used as a preamble to further population genetic analysis. Char were analyzed using STRUCTURE but synthesis and inferences were not completed in time to meet the reporting deadline. Note that Cutthroat Trout controls and hybrid *O. mykiss* (as determined by three diagnostic SNPs) were not removed from dataset analyzed using STRUCTURE to qualitatively explore the effect of undetected hybridization on population structure. STRUCTURE performed individual analyses for an assumed population number (K) from two to ten. For each K, three independent runs were conducted, each having a 100,000 MCMC iterations burn-in period followed by population estimation consisting of 100,000 MCMC iterations. The admixture ancestry model was used, and allele frequencies were assumed to be correlated among populations. STRUCTURE HARVESTER (Earl and VonHoldt 2012) was used

to aggregate and summarize the results output. Evaluating the K that best fit the data followed the Evanno method, as implemented in STRUCTURE HARVESTER.

4.4.5 Year 2 Effective Population Size

The program LDNE (Waples and Do 2008) was used to estimate N_e and N_b of Bull Trout, Dolly Varden, and Rainbow Trout in the study area. The program assumes samples are collected randomly from a well-defined population that does not receive migration and that markers are unlinked, neutral, and in Hardy-Weinberg proportions. The software uses LD (nonrandom association of alleles among loci) to estimate the inbreeding-effective size, which is the size of an ideal population (i.e., at HWE) that would result in the same reduction in heterozygosity as in the actual population being considered. When the method is applied to samples from individuals that are the same age (cohorts), LDNE estimates N_b within the year prior to that sampled. The presence of rare alleles in dataset tends to upwardly bias effective size estimates (Waples and Do 2008). To provide a balance between precision and bias, alleles that were rare (frequencies < 0.02) were ignored. An additional bias correction of Waples et al. (2014) for overlapping generations was applied, which is based on estimates of adult life span and age of first reproduction obtained from the peer-reviewed scientific literature.

4.4.6 Year 2 Age Determination for Effective Size Estimation

4.4.6.1 Scale Analysis

The age of each sample was estimated to inform estimates of N_e and N_b in the study area. Scale collection, preparation, and aging followed the general protocols of Love (2016) and Copeland et al. (2018). Samples were dried and the three most legible scales from each individual were mounted between two microscope slides for imaging. High-resolution scale images were captured using a microscope, a digital microscope camera and Image Pro Premier version 9.2 software.

A subset of individual scales selected based on size class from each genus were visually analyzed for age based on the number of annuli. Un-aged scales were assigned ages (see Section 5.3.1.4). Analysis was completed at the genus level because genetic identification were not available yet and sample sizes were small in Bull Trout. Discussion of potential bias of completing the analysis at the genus level is in the Section 6.0 of this report. Salmonid scales are difficult to age due to their variable life history and individual differences in scale reabsorption during stressful periods, leading to misidentification of annuli (Hernandez et al. 2014). Independent age estimates were performed by two reviewers. Age determinations for each individual were compared between reviewers; if an age difference occurred between the reviewers, a more senior reviewer resolved the difference. Due to the uncertainty in aging, all age classes are reported as X+ indicating that the age is at least that old, but the exact age is undetermined.

4.4.6.2 Length-at-Age Key

A length-at-age key based on the subset of scale aged fish and their fork length was constructed for each genus using the R package {FSA} (Ogle et al. 2022). Fish of unknown ages were assigned to age using the semi-random method (Isermann and Knight 2005) based on the length-at-age key. Fish with lengths outside of the range of the length-at-age key fish were not assigned an age.

4.4.7 Year 2 Haplotype Diversity Associated with Life-history

No life history data was collected in Year 2 and so any inference with respect to adaptive genetic diversity is speculative. For instance, it is generally accepted that genetic variation associated with adaptive traits, like juvenile propensity to migrate, is inconsistent across the range of *O. mykiss*. However, the genotyping panel used in this study included loci that are potentially adaptive (i.e., not neutral), and LPs have asked for basic summary statistics for some markers. Specifically, the panel includes markers for genetic regions indicative of juvenile emigration propensity (frequency of OMY5), adult return timing to freshwater (frequency of OMY28), and adaptive diversity associated with climate, land cover, stream temperature, elevation, wind velocity, solar radiation, and stream network variables. The RSP did not specify any analysis needs regarding adaptive diversity, so these data will remain largely unevaluated. Yet, the study team should note that questions submitted by LPs in Fall 2021 and discussed by the Expert Panel in 2022 prior to Year 2 activities stated an interest in evaluating life history differences. A cursory evaluation regarding adaptive diversity associated with juvenile propensity to emigrate (chromosome 5 loci; OMY5) was conducted to demonstrate estimation of haplotypes present at putatively adaptive loci data generated in this study. In the Reservoir Fish Genetics Study, the exercise was to document the different haplotypes observed in the study area and their base condition frequencies. It is important to note that just because a genetic marker is associated with an adaptive trait in one population, it does not necessarily mean that the same marker must be associated with the same trait in all populations. This is because genetic associations can vary depending on a variety of factors, including environmental and genetic reasons.

For four OMY5 loci present in the genotyping panel, Pearse et al. (2014) determined for *O. mykiss* from California that a specific haplotype was associated with *O. mykiss* juveniles emigrating downstream (exhibiting anadromous behavior) and a second haplotype was associated with freshwater residency (non-migrating). While the correlation between OMY5 haplotypes and juvenile behavior published by Pearse et al. (2014) was well founded in the southern range of Rainbow Trout, the haplotype relationships may not apply in the Study Area. The emigration-associated haplotype would appear in this study for locus OmyR14589, OmyR19198, OmyR24370, and OmyR33562 as 4-3-3-1, for a T, G, G, and A collocated along the same chromosome within an individual. The residency haplotype would be a 1-1-1-3. Haplotypes present in the study area and their frequencies were determined using {haplo.stats} package in R (Sinnwell et al. 2022). Global frequencies across all individuals were estimated as well as frequencies within each population.

5.0 RESULTS

5.1 Expert Panel

The Reservoir Fish Genetics Study was conducted in consultation with an advisory Expert Panel. In accordance with the RSP (City Light 2021) and the June 9, 2021 Notice, City Light convened an Expert Panel composed of resource agency specialists and experts from academia with backgrounds in genetics and/or ecology. The purpose of the Expert Panel was to provide input and recommendations to inform City Light's study approach and decisions at specific milestones. Three meetings were to be held with the Expert Panel throughout the study process; in addition, one working session was held by the Expert Panel independently, as well as one Expert Panel-LP coordination call. Expert Panel meetings are summarized below and provided in Attachment D.

The members of the Expert Panel included:

- Hope Draheim (USFWS);
- Jason Dunham (USGS);
- Alex Fraik (NMFS Affiliate);
- Jim Meyers (NMFS);
- Meryl Mims (Virginia Tech);
- Krista Nichols (NMFS);
- Carl Ostberg (USGS);
- George Pess (NMFS);
- Todd Seamons (WDFW);
- Matt Smith (USFWS);
- Adrian Spidel (NW Indian Fisheries Commission); and
- Rick Taylor (University of British Columbia).

5.1.1 Expert Panel Meeting No. 1

On October 26, 2021, the first Expert Panel meeting was held with Expert Panel members and LPs as a study kick off and "meet and greet" opportunity with the experts. Expert Panel members, LPs, the consultant team (including CFS), and City Light were in attendance. Erin Lowery (Seattle City Light) provided a brief overview of the Reservoir Fish Genetics Study and the goals of the Expert Panel, milestones, and timelines. E. Lowery emphasized the critical need for the Expert Panel and their advisory role on establishing native fish population baseline information to support the identification of potential future management objectives in the Upper Basin. The meeting also included a "question and answer" session with the Expert Panel. The scope of the Expert Panel's role, the Year 1 Existing Data Review Technical Memorandum (Attachment A), and action items were also discussed.

5.1.2 Expert Panel Meeting No. 2

The second Expert Panel meeting was held on January 18, 2022. The purpose of the meeting was to discuss the information contained in the Year 1 Existing Data Review Technical Memorandum (Attachment A) and the research questions proposed by LPs. Expert Panel members, LPs, the consultant team (including Cramer Fish Sciences), and City Light were in attendance. Dan Bingham (CFS) provided an overview of the existing genetics information contained in the technical memorandum and answered questions from LPs and the Expert Panel. Discussion topics included the use of microsatellite data versus SNPs, Dolly Varden and Bull Trout population divergence above and below Gorge Dam, and availability of metadata and additional tissue samples. Dan Bingham also explained the framework through which the study team would be analyzing and answering LP and City Light questions and objectives related to the genetic management of native fishes. He reviewed the study's objectives as outlined in the RSP and described the congruence between City Light's objectives and LP questions by examining potential metrics to analyze and answer LP questions. The Expert Panel discussed the metrics to be used in the study and their sufficiency in answering the questions proposed by the LPs. There was broad discussion and agreement on the need for the Expert Panel, LPs and City Light to further define how genetic monitoring could support long term reservoir fish management decisions.

5.1.3 Expert Panel Working Session

The Expert Panel held an independent working session on February 9, 2022 to discuss the data, approach, questions, and objectives of the Reservoir Fish Genetics Study. T. Seamons and M. Smith provided background on the existing information and prior research. The group discussed the need for LPs to refine their study questions and needs from City Light/the consultant team in order to make informed advisement on the process and study.

5.1.4 Expert Panel Coordination Meeting

On March 31, 2022, a meeting with the Expert Panel and LPs was held to facilitate communication between the parties with an objective to further refine study questions submitted by LPs in December 2021. The group walked through the LP questions, examining the overview section and specific questions in greater detail. They identified aspects of the document and questions that required greater context, detail, and relevancy.

5.1.5 Expert Panel Meeting No. 3

The final meeting was held on January 30, 2023, to present the results of the two-year study and discuss and identify relevant topics and/or management objectives for consideration in long-term reservoir fish management at the Project. A technical memorandum was provided to the Expert Panel in advance of the meeting. The technical memo provided an overview of the study results to support meeting objectives. Information in the technical memo was similar but not identical to information presented in this study report. Following the meeting, City Light requested that written comments from Expert Panel members be submitted, with an emphasis on applied management topics related to genetics, i.e., species conservation, connectivity, stock identification, and potential future Skagit reservoir fish management. Several Expert Panel members provided comments for consideration. As appropriate, comments relevant to study objectives provided by Expert Panel members were incorporated in this study report, as the purpose of the technical memo was only to provide preliminary results of the study to the Expert Panel to support discussion at

the January 2023 meeting. The technical memo is provided as Attachment C, and the meeting summary provided in Attachment D. Note that the meeting summary was not finalized until after filing of the USR, therefore the summary attached is preliminary.

5.2 Year 1 Results

5.2.1 Year 1 Rainbow Trout (*Oncorhynchus mykiss*)

5.2.1.1 Year 1 Collections (existing data)

In the data provided by WDFW in July 2021 (described in Section 4.1 of this study report), the microsatellites appeared to be a subset of the standardized Stevan Phelps Allele Nomenclature (SPAN) markers described in Stephenson et al. (2009) that were developed to ensure data quality (repeatable allele scoring) across laboratories. The data were provided in a Microsoft Excel spreadsheet that included Sample ID, WDFW Collection Code, Count, and Percent Missing Data. Exact sampling locations were not provided but collections appeared to be from the sites in the Skagit and Fraser River basins that are described in Pflug et al. (2013), which included tributaries, mainstem rivers, hatcheries, and Project reservoirs. Some sites appeared to have been sampled across multiple years. The collections ranged in size from 1 in the Suiattle River in 2009 to 106 in Diablo Lake in 2005. No metadata were provided regarding sampling field methods (e.g. electrofishing), whether samples were collected randomly, or targeted life stages, life histories, morphologies, taxa, etc.

Of the 2,697 samples provided by WDFW, 536 were removed due to missing genotypes at two or more loci (e.g., Reeves et al. 2016), and 20 were removed because of duplicated genotypes. Pooling of samples from the same locations across years reduced the number of analyzed collections from 76 to 25; however, only four of the pooled collections were retained due to decreased deviation from HWE: Bacon Creek (2007 to 2010), Clear Creek (2009 and 2010), Blackwater River (2009 and 2010), and the Suiattle River (1981 and 2009 to 2011). Putative siblings from the *O. mykiss* dataset were not omitted because multiple age classes appeared to have been sampled and doing so could reduce precision of analyses as cautioned by Waples and Anderson (2017) (but see analyses of native char). Collections with fewer than 25 individuals were removed to avoid biased estimates of allele frequencies within sub-populations (Hale et al. 2012). Data were not sufficient to describe hybridization with Cutthroat Trout (*O. clarkii*) because the submitted spreadsheet from WDFW did not contain known nonhybridized Cutthroat Trout genotypes to use as positive controls for estimating taxon-diagnostic allele size distributions. The final dataset contained 1,900 individuals from 40 collection events but only 38 collections were analyzed due to possible hybridization with *O. clarkii* that was not apparent until most analyses were completed. The genotypes are available upon request in GENEPOP format (Rousset 2020).

5.2.1.2 Year 1 Genetic Summary Statistics

For previously reported microsatellite data, comparison of observed (H_o=0.729) and expected (H_S=0.747) heterozygosity across all collections and loci suggested a relatively small but overall deficit of heterozygotes (F_{IS}=0.025, 95 percent confidence interval [CI]; 0.01, 0.03). As described in the Preamble to the Methods, comparison of H_o and H_S is a fundamental way of determining whether a collection represents a single population. Eighty-six of 600 (14 percent) randomization tests for HWE (15 markers x 40 collections) using FSTAT (Goudet 1995) were significant at the α=0.05 level with 68 (79 percent) of the tests showing a deficit of heterozygotes. No tests for HWE

were significant at the adjusted level of α=0.00008. The locus *One-14* deviated from HWE in 17 of 40 (42.50 percent) total collections with all tests showing a deficit of heterozygotes. By contrast, most other markers (11 of 15) produced various combinations of heterozygote excess and deficiency. Therefore, the locus *One-14* was omitted from further analysis due to the possibility of genotyping problems. This adjustment decreased mean F_{IS} to 0.017, though the difference was not statistically significant (95 percent CI: 0.03, 0.01). The remaining 14 microsatellite loci had a total of 312 alleles, ranging from 11 at *Ots-4* to 32 at *Omy-1001*. Across all 14 loci and 40 collections, the estimated false detection rate of a parent-offspring pair was 0.00000811, 0.00000033 of full siblings, and 7.277×10^{-21} of unrelated individuals. However, within any single collection, power is expected to be substantially lower. For example, the false positive rate (FPR) for related individuals in Roland Creek, a tributary to Ross Lake, is 0.0000161 and the false negative rate (FNR) is 0.392, a result of assignment posterior probability overlap given individuals of known relationship (Figure 5.2-1). Expected heterozygosity (H_S) within each collection ranged from 0.36 in the collection from North Fork Cascade River in 2010 to 0.83 in the Baker River in 2010 (Table 5.2-1). Average gene diversity in collections from Gorge Lake (H_S=0.74) was the same as the diversity in all other collections (H_S=0.74).

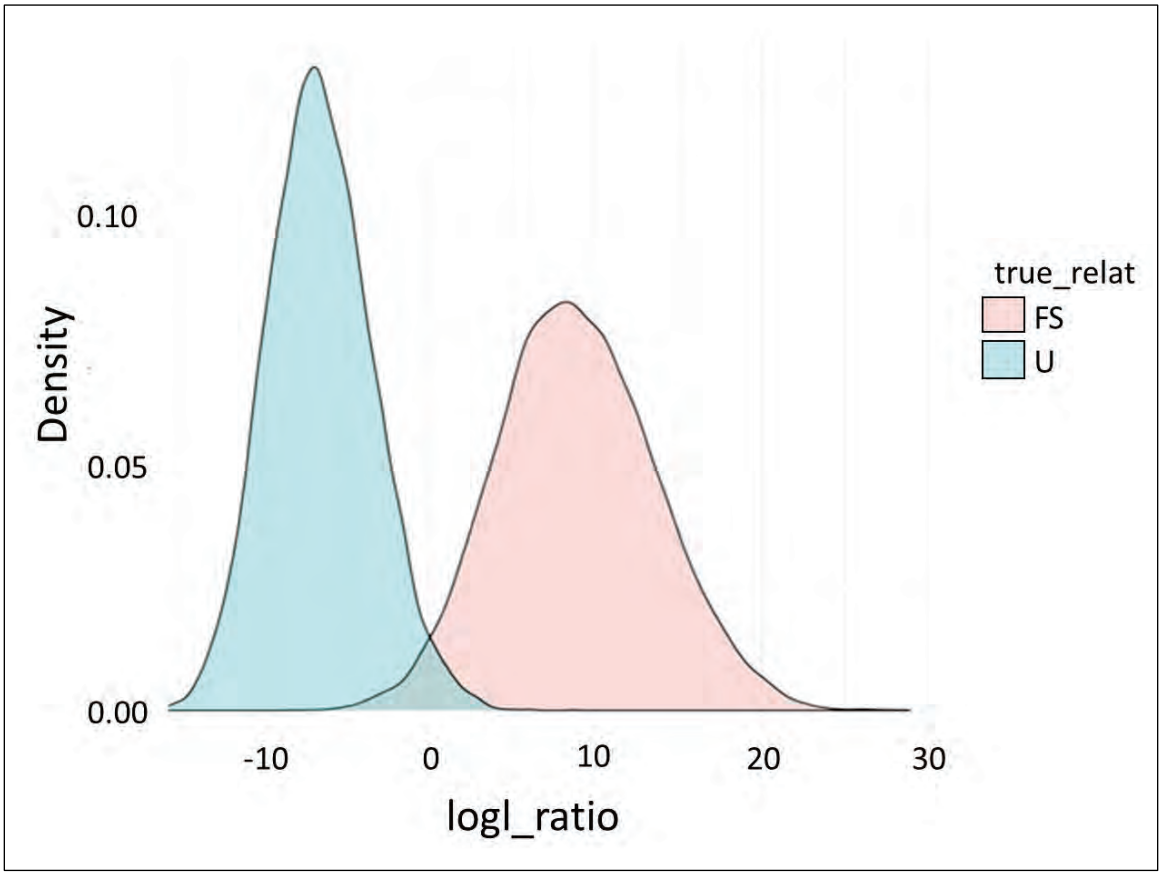


Figure 5.2-1. **Log-likelihood ratio distribution for simulated true full-siblings versus unrelated individuals based on Roland Creek *O. mykiss* genotype data from Year 1. The analysis used genotypes from 15 microsatellites that were provided by WDFW. Legend: FS=Full Sibling; U=Unrelated. High overlap between full-siblings and unrelated fish suggests relatively low statistical power to detect highly related individuals.**

Table 5.2-1. Year 1 summary statistics for samples collected from *O. mykiss* in the Skagit and Fraser River basins. Analysis of these data have been previously reported, for example, in Pflug et al. (2013). Genotypes consisted of 15 microsatellite loci. The collections were provided by WDFW. The collections came from various places, including the Project area, downstream of the Project area, and adjacent watersheds. Some collections were from the anadromous zone (identified in the table notes).

Collection number	Collection size	WDFW Code [1]	Location	Origin[2]	Upper Skagit[3]	Stage	Phenotype[4]	F_{IS} [5]	H_S [6]	A_R [7]	R^2 [8]
1	57	07MS, 08MI, 10BA	Bacon Creek	NOR	No	Juvenile, adult		0.01	0.79	9.45	0.02
X	57	09EL	Baker River 09	NOR	No		Trout	0.09	0.82	10.44	0.03
X	42	10AU	Baker River 10	NOR	No		Trout	0.11	0.84	11.13	0.04
2	51	09EU	Big Creek 09	NOR	No		Trout	0.04	0.66	5.63	0.02
3	48	10BG	Big Creek 10	NOR	No		Trout	0.06	0.67	5.16	0.03
4	52	09JB, 10BJ	Blackwater River*	NOR	No	Juvenile	Trout	0.11	0.74	7.61	0.02
5	66	10MZ	Chilliwack Hatchery	HOR	No	Adult		0.00	0.76	8.02	0.02
6	94	09ET, 10BE	Clear Creek	NOR	No		Trout	0.06	0.68	8.56	0.01
7	38	10BB	County Line Ponds	NOR	NO	Juvenile		0.01	0.80	9.10	0.04
8	26	05NG	Diablo*	NOR	Yes		Trout	0.06	0.75	8.50	0.04
9	41	10BK	Diobsud	NOR	No	Juvenile		0.02	0.79	9.77	0.03
10	43	03OA	Dry Creek	NOR	Yes		Trout	0.02	0.71	7.50	0.03
11	47	09EH	Finney Creek*	NOR	No	Juvenile		0.00	0.78	9.27	0.03
12	47	10AT	Finney Creek*	NOR	No	Juvenile		0.01	0.80	9.59	0.02
13	30	11BK	Finney Creek*	NOR	No	Adult		-0.02	0.80	10.40	0.04
14	38	09IZ	Goodell Creek	NOR	No	Juvenile		0.01	0.77	8.76	0.03
15	41	10BC	Goodell Creek	NOR	No	Juvenile		0.00	0.79	9.12	0.03
16	47	09EE	lower Cascade	NOR	No	Juvenile		-0.05	0.77	8.23	0.03
17	44	10AV	lower Cascade	NOR	No	Juvenile		0.03	0.79	9.26	0.03
18	48	10AY	lower Skagit	NOR	No	Juvenile		0.02	0.79	9.51	0.03
19	28	08LF	lower Skagit	NOR	No	Adult		0.01	0.78	9.26	0.04
20	59	09CF	Marblemount	HOR	No	Adult		0.01	0.82	9.68	0.02
21	44	10AN	Marblemount	HOR	No	Adult		0.03	0.79	8.89	0.03

Collection number	Collection size	WDFW Code [1]	Location	Origin[2]	Upper Skagit[3]	Stage	Phenotype[4]	F_{IS} [5]	H_S [6]	A_R [7]	R^2 [8]
22	39	09BM	mid Skagit	NOR	No	Adult		0.01	0.80	10.49	0.03
23	31	10AS	mid Skagit	NOR	No	Adult		0.04	0.80	10.14	0.03
24	47	09ES	NF Cascade*	NOR	No		Trout	0.11	0.41	4.30	0.02
25	45	10BF	NF Cascade*	NOR	No		Trout	-0.08	0.36	3.98	0.02
26	79	02FB	Roland Creek*	NOR	Yes		Trout	0.01	0.71	7.68	0.01
27	30	06AF	Ross*	NOR	Yes		Trout	0.03	0.73	8.20	0.04
28	44	09MA	Ross*	NOR	Yes		Trout	-0.01	0.69	6.65	0.04
29	47	10BH	Ross*	NOR	Yes		Trout	-0.03	0.70	6.40	0.04
30	45	10AX	Sauk*	NOR	No	Juvenile		0.04	0.80	9.66	0.03
31	29	83AAA	Sauk*	NOR	No	Adult		0.06	0.80	10.29	0.04
32	32	09JA	Stetattle*	NOR	Yes		Trout	0.03	0.76	8.66	0.04
33	41	10BI	Stetattle*	NOR	Yes		Trout	0.03	0.77	8.79	0.03
34	115	09DT, 09EF, 10AQ, 10AW, 11BM	Suiattle*	NOR	No	Juvenile, adult		0.01	0.79	10.05	0.01
35	51	09EV	upper Finney*	NOR	No		Trout	0.03	0.74	6.52	0.02
36	49	10BD	upper Finney*	NOR	No		Trout	0.04	0.72	6.77	0.02
37	56	10AZ	upper Skagit	NOR	No	Juvenile		0.01	0.79	9.56	0.02
38	32	11BI	upper Skagit	NOR	No	Adult		0.00	0.81	10.43	0.03

1 WDFW code: WDFW collection identification with apparent sample year as the prefix.
2 Origin: hatchery (HOR) or natural (NOR) origin.
3 Upper Skagit: collections from upstream of Gorge Dam in the Skagit River and from British Columbia.
4 Phenotype: identifies whether collections were from apparent trout as determined by WDFW.
5 F_{IS}: estimated deviation from HWE. Positive result means there was a deficit of heterozygotes, whereas negative value means there was an excess of heterozygotes. Heterozygote excess is expected in small populations due to random allele frequency differences between males and females.
6 H_S: estimated expected heterozygosity within sub-populations (i.e., gene diversity).
7 A_R: estimated allelic richness.
8 R^2: is the estimated pairwise correlation of alleles among loci.
9 * Collections of resident *O. mykiss* from upstream of anadromous barriers include Finney Creek, Clear Creek (Upper Sauk basin), Big Creek (Upper Suiattle River) , North Fork Cascade River and collections upstream of Gorge Dam.

Six-hundred-forty of 3,640 (17.5 percent) log-likelihood (*G*) tests for pairwise LD using FSTAT were significant at the α=0.05 level. However, only 15 (less than 1 percent) tests were significant at the adjusted table-wide level of α=0.00007. As described in the Methods Preamble, there is a difference in alpha/significance level between the tests for HWE (α=0.00008) and LD (α=0.00007) because HWE and LD are two different "families" of tests (i.e., 600 tests for HWE and 3,640 tests for LD, respectively). Thus, the FWER is different between them. The greatest disequilibrium was observed in the collection from Diablo Lake in 2005 (R^2=0.04) and the least in Suiattle River (R^2=0.01) (Table 5-2.1).

High levels of LD can have multiple causes, including sampling, demographic, and evolutionary factors. From a sampling perspective, high LD can result from unwittingly merging two populations into one sample or from overrepresenting certain families in a sample. In terms of demographics and evolution, high LD can occur through hybridization, recent common ancestry, population bottlenecks, small effective population size, non-random mating, and selection, among other factors. If the samples are collected from well-defined populations that are randomly mating and receiving no migration (gene flow), high LD is likely an indicator of effective population size. Thus, the high levels of LD observed in the Diablo Lake sample could be due to the aggregated nature of the sample, which might contain fish from multiple populations (i.e., because it appears to be an at large collection from the reservoir as opposed to a collection from a tributary).

Fisher's exact tests using POWSIM (Ryman and Palm 2006) which were based on sample sizes and estimated allele frequencies of the dataset, suggested power to detect deviation from genetic homogeneity was 0.32 for F_{ST}=0.001 and was 1.00 for F_{ST}=0.01. The overall estimated proportion of genetic variance explained by population structure (F_{ST}) was 0.094. Log-likelihood (*G*) tests for population differentiation were significant for each locus and across all loci (P less than 0.001). Estimates of pairwise F_{ST} ranged from -0.004 between collections from Stetattle Creek in 2009 and 2010 to 0.39 between collections from Ross Lake and North Fork Cascade River.

Principal Component Analysis (PCA) of individual-based genetic distances using {adegenet} (Jombart 2011) accounted for a relatively small amount of projected inertia—a metric of the magnitude of the explained genetic variance among individuals (cumulative inertia explained by PC-1 through 3=5.924 percent). In general, if the first few PCs explain a small amount of the inertia, it suggests that they are not capturing much of the variability in the data and that further exploration or alternative analysis methods may be necessary to gain insights into the structure of the data. Nevertheless, genetic population structuring was apparent in scatterplots of the first three PC. However, several samples from the Baker River collections appeared to be outliers along axes 1 and 2. Notes provided by WDFW suggested the samples could be hybrids with *O. clarkii*. Reanalysis without the Baker River collections only slightly improved projected inertia of the first three PCs (6.095 percent); however, it did improve visualization of genetic population structure (Figure 5.2-3). Specifically, PC-1 (2.215 percent) clearly distinguished the North Fork Cascade River (Collections 24 and 25) from all other collections. PC-2 (2.044 percent) highlighted additional population structuring with collections from the study area tending to display positive inertia, collections from the Sauk River basin tending to display negative inertia, and remaining collections falling in between. PC-3 (1.836 percent) nearly distinguished Big Creek (Collections 2 and 3) from all other collections (Figure 5.2-4).

Limiting PCA to collections from study area at Gorge Lake identified three samples that might be hybrids between *O. mykiss* and *O. clarkii* based on notes from WDFW; they were subsequently removed from the analysis (09JA0030, 05NG0056, and 10BI0047). Reanalysis without the potential hybrid samples indicated that the first three PCs explained 5.898 percent of the total inertia (Figure 5.2-5) and appeared to support some genetic structuring associated with location but statistical support for individual genetic groups was low.

In Year 1, effective population size (N_e) of *O. mykiss* was not estimated in the Project reservoirs because of existing data limitations. Firstly, presence of hybrids within dataset was unclear and hybridization with *O. clarkii* could bias estimates of N_e by creating genetic disequilibria that is not associated with genetic drift. Secondly, estimating N_e in an iteroparous species with overlapping generations requires extensive sampling effort and significant data on life stage specific survival and reproduction. Though it is common to estimate effective number of breeders (N_b), unbiased estimates typically call for sampling of individuals of the same cohort or across multiple generations, and such data in Year 1 was not available. An objective in Year 2 was to estimate N_e / N_b for Study Area populations where possible.

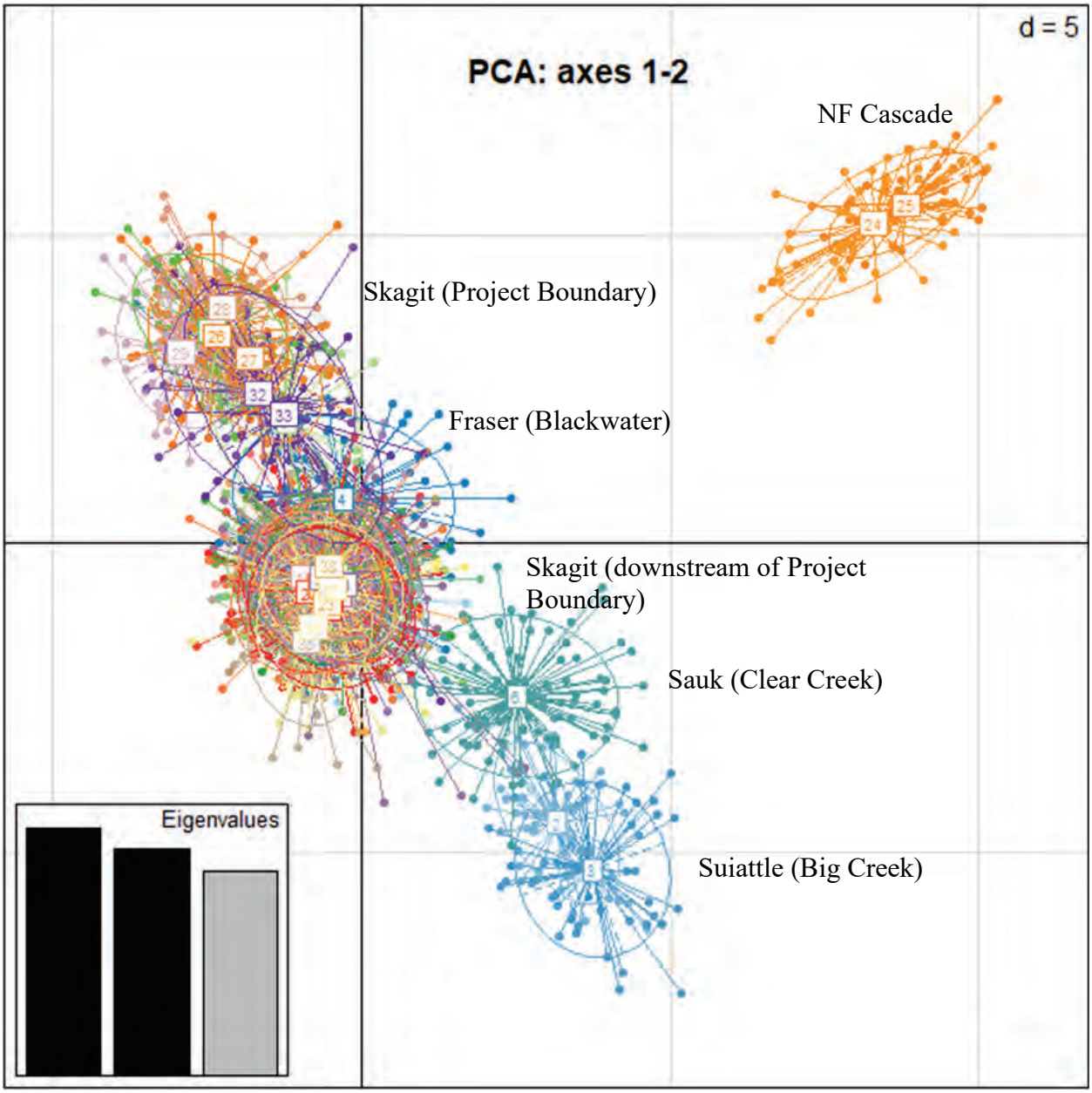


Figure 5.2-2. **Scatterplot of PC-1 (2.215 percent) and PC-2 (2.044 percent) for Year 1 Rainbow Trout microsatellite data for all collections, excluding samples from the Baker River, estimated using {adegenet} in program R (Jombart 2011). The distribution of genetic variation appears to support some genetic structuring associated with the geographical locations of collections. River basins are provided to show the approximate geographical locations of each collection. NF Cascade was apparently collected upstream of a barrier. Metadata and types of collections (i.e., resident/adult) are shown in Table 5.2-1. Numbers at centroids identify the collection number listed in Table 5.2-1. Ellipses define 1.5 standard deviations of the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom left corner shows first three eigenvalues.**

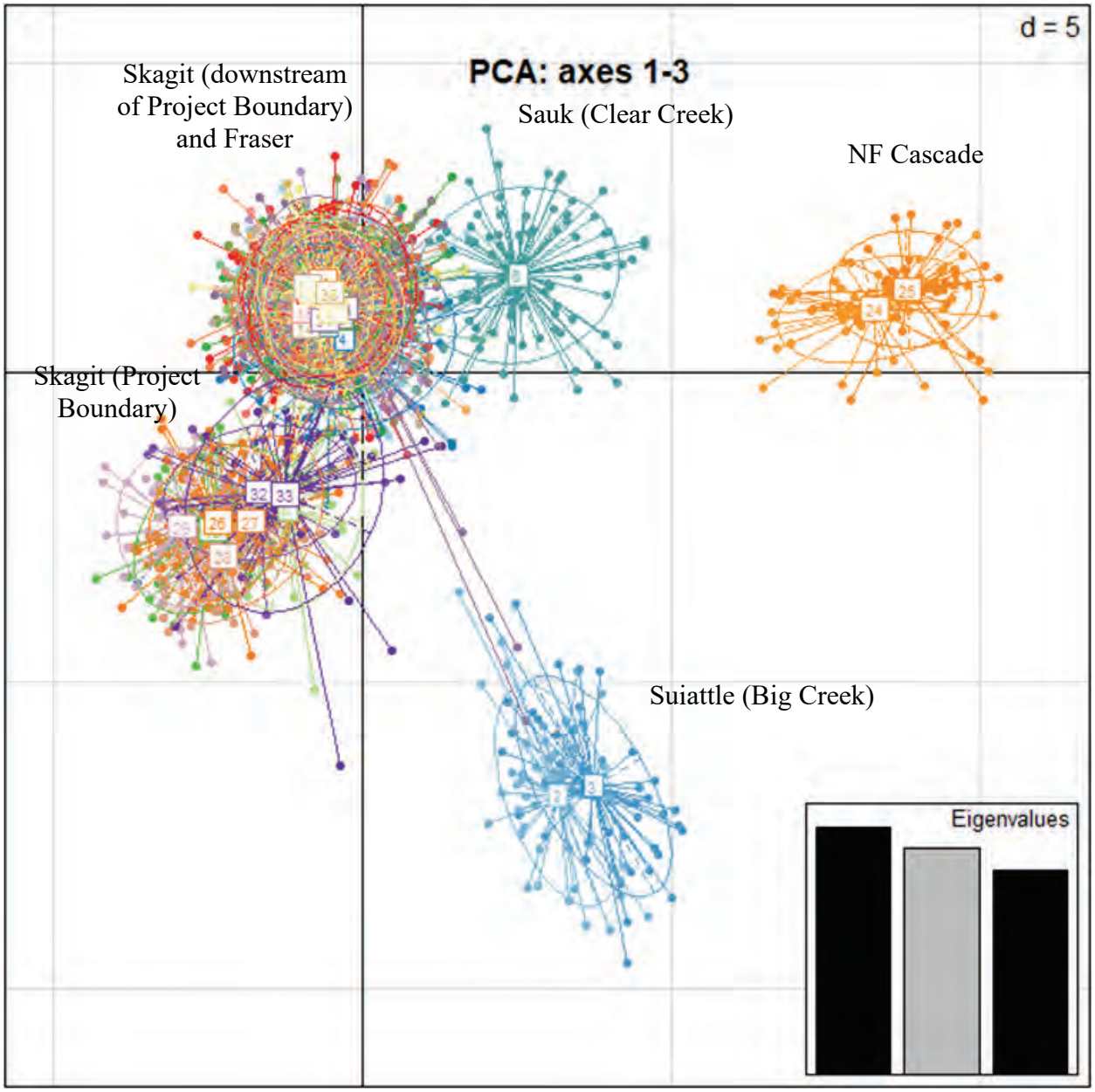


Figure 5.2-3. **Scatterplot of genetic PC-1 (2.215 percent) and PC-3 (1.836 percent) for all Rainbow Trout collections, excluding samples from the Baker River, estimated using {adegenet} in program R. The distribution of genetic variation appears to support existence of genetic structuring associated with the geographical locations of collections. River basin names are provided to describe the approximate geographical locations of each collection. Numbers at centroids identify the collection number listed in Table 5.2-1. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom right corner shows first three eigenvalues.**

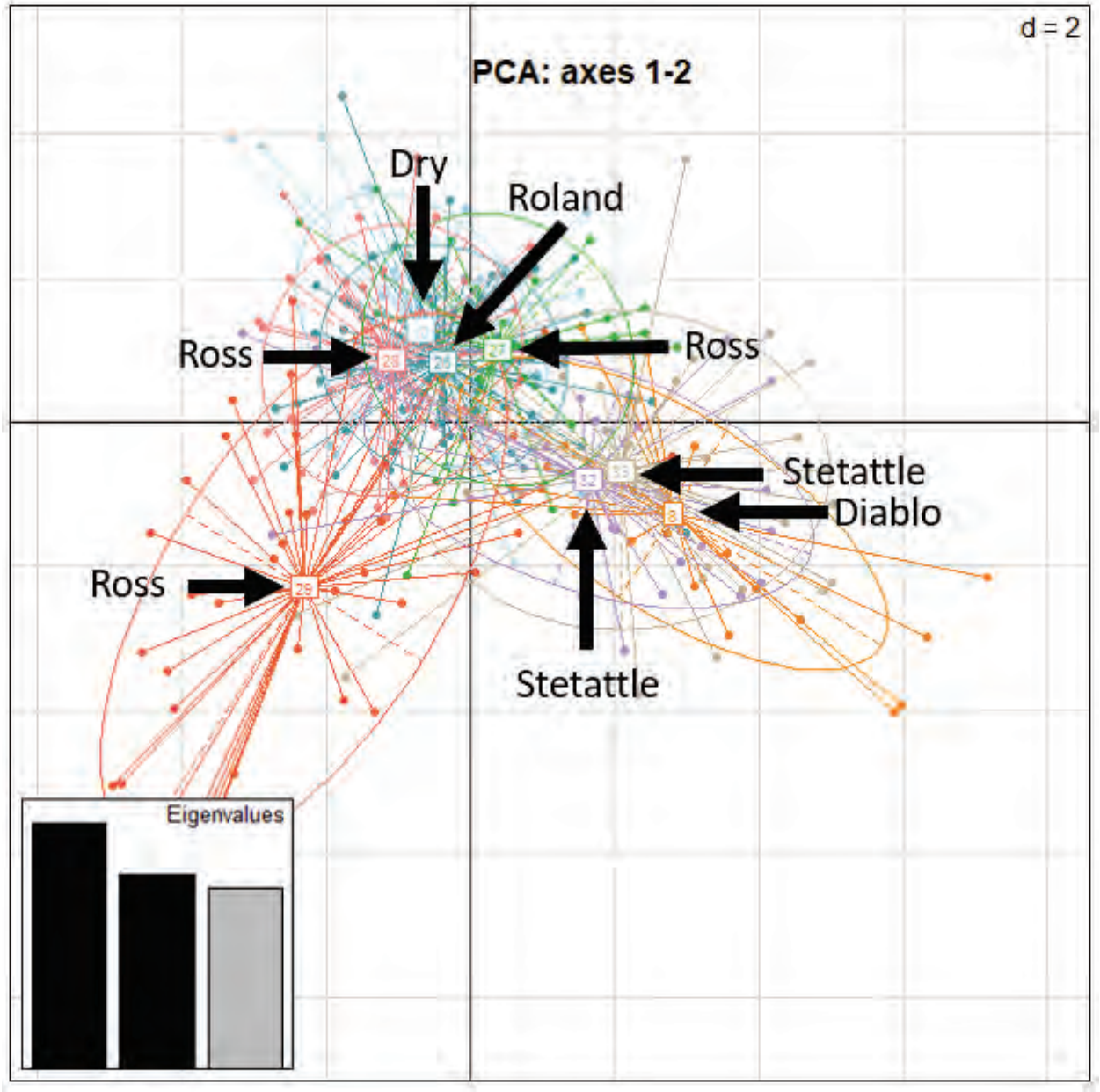


Figure 5.2-4. **Scatterplot of genetic PC-1 (explaining 3.870 percent of the variation) and PC-2 (explaining 2.028 percent of the variation) for all Rainbow Trout collections located upstream of the Gorge Dam. The plot was generated using the {adegenet} package in the R programming language and 15 microsatellites from Year 1 were used. The numbers at the centroids correspond to the collection number listed in Table 5.2-1. The scree plot in the bottom right corner displays the first three eigenvalues. The ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more indicate less distinction between groups. The scree plot in the bottom left corner also shows the first three eigenvalues. It is important to note that inertia is a measure of relative genetic differences between individuals (as outlined in the Preamble of the Methods section), and should not be confused with F_{ST}, which measures divergence at the population level.**

5.2.2 Year 1 Native Char (*Salvelinus* spp.)

5.2.2.1 Year 1 Collections (existing data)

Eight hundred and ninety-eight *Salvelinus* spp. genotypes at 16 microsatellite loci were provided by USFWS and WDFW following a request for existing Bull Trout data within the Skagit River basin (Table 5.2-2). USFWS provided 563 of the genotypes and WDFW provided 335. The standardized markers included *Omm1128*, *Omm1130* (Rexroad et al. 2001); *Sco102*, *Sco105*, *Sco106*, *Sco107*, *Sco109* (WDFW unpublished); *Sco200*, *Sco202*, *Sco212*, *Sco215*, *Sco216*, *Sco218*, *Sco220* (Dehaan and Ardren 2005); *Sfo18* (Angers and Bernachez 1996); and *Smm22* (Crane et al. 2004). The collections were from four study area tributaries (upper Skagit River, Big Beaver, Ruby, and Stetattle creeks) and all three reservoirs (Ross, Diablo, and Gorge lakes). It was unclear which *Salvelinus* spp. taxa or their hybrids were included in the dataset. It was also unclear to what extent collections comprised highly related individuals, which is a common concern in genetic studies of Bull Trout (DeHaan et al. 2014). Furthermore, USFWS communicated that the juvenile collections likely contained related individuals (Smith 2021).

Sampling location metadata were not provided for USFWS samples, so sampling locations were assumed to be the same as reported in Smith (2010). The stated purpose of the collections from Smith (2010) was to assess genetic variability within and between Bull Trout populations, with sampling methods including a combination of electrofishing, snorkeling, and angling.

No metadata were provided by WDFW other than collection code. Location data were not provided, so samples obtained from within the study area were considered "at-large" from reservoirs. The stated purpose of WDFW collections was to characterize the genetic variation of Bull Trout, Dolly Varden, and Brook Trout in the Project reservoirs, but no collection methodology was described. The degree to which samples were collected randomly across *Salvelinus* spp. taxa was unknown, including whether any special effort was made to target Bull Trout, Dolly Varden, or Brook Trout or whether potential hybrids were targeted or avoided. Sampling considerations are a key concern because targeted collections (i.e., based on morphology) can bias inference into studies of genetic variation.

Quality assurance/quality control procedures were performed to obtain a final dataset in which basic population genetic analyses could be reasonably implemented. Duplicate genotypes were observed for sample IDs 12FG008 and 12FG0009, and so sample 12FG0009 was omitted from the dataset. All individuals with missing genotypes at three or more loci were removed. This lower threshold for removing fish incomplete genotypes was necessary, however, because the Bull Trout data production appeared to have been conducted in four by four-locus panels (i.e., multiplexes), with many samples missing a single four locus block. Following data quality assurance/quality control, 589 samples were successfully genotyped and retained for evaluation (Table 5.2-2), although individuals suspected as being hybrids were subsequently removed before analysis.

Table 5.2-2 Summary of Year 1 Bull Trout microsatellite dataset collection provided by WDFW (335) and USFWS (563). These samples were analyzed using 16 microsatellites and 589 were retained and evaluated. Suspected hybrids were removed before analysis.

Collection Location	River	Life Stage	Year (WDFW Code)	Number Collected	Number Evaluated	Number Analyzed
Upper Skagit River	Skagit	Adult	2001	16	14	14
Big Beaver Creek	Skagit	Adult	2009	21	21	21
Ruby Creek	Skagit	Adult	2001, 02, 04, 09	43	41	41
Stetattle Creek	Skagit	Juvenile	2009	59	41	41
Lower Goodell Creek	Skagit	Juvenile	2009	60	46	46
Upper Goodell Creek	Skagit	Juvenile	2009	19	8	8
Bacon Creek	Skagit	Juvenile	2009	61	24	24
Cascade River	Cascade	Juvenile	2009	39	33	33
Marble Creek	Cascade	Juvenile	Unknown	28	18	18
Kindy Creek	Cascade	Juvenile	Unknown	30	17	17
Illabot Creek	Sauk	Juvenile	2009	70	60	60
South Fork Sauk River	Sauk	Juvenile	2009	59	54	54
Downey Creek	Sauk	Juvenile	Unknown	58	44	44
Ross Lake	Skagit	unknown	2012 (12FG)	54	47	42
Ross Lake	Skagit	unknown	2015 (15OW)	28	22	20
Diablo Lake	Skagit	unknown	2013 (13PS)	40	29	8
Gorge Lake	Skagit	unknown	2014 (14ST)	27	5	3
Gorge Lake	Skagit	unknown	2019 (19NL)	109	22	0
Sulfur	Skagit	unknown	2005 (050F)	4	4	4
Sulfur	Skagit	unknown	2006 (06JQ)	28	23	23
Diablo, Gorge Lake	Skagit	unknown	2011 (11LX)	45	16	9
Total				**898**	**589**	**530**

5.2.2.2 Year 1 Identification of Related Individuals within Collections

Statistical power was estimated to correctly classify related individuals. This was completed to evaluate the possible effects of violations of sampling assumptions common to the analysis of Bull Trout microsatellite data; specifically, that highly related individuals (i.e., full siblings) are common in samples of Bull Trout (particularly samples of juveniles), which can result in pseudo-replication of genotypes and thus biased estimates of allele frequencies (DeHaan et al. 2014). Statistical power of pedigree analysis to identify parent-offspring and full-sibling pairs was conducted using the R package {CKMRSim} version 0.1 (Anderson 2019). During pedigree analysis, all samples are examined for relatedness in pairwise comparisons, and so the FPR increases exponentially with sample size. It is recommended to choose a FPR threshold approximately 10 times smaller than the reciprocal number of pairwise comparisons. In this case,

1.4 e^{-5} was the target FPR used to evaluate the power to detect relatives (i.e., 0.10 * (100 x 100)$^{-1}$ = 0.000014). To simulate the related and unrelated individuals needed to estimate power of pedigree analysis, all collections from the Skagit River dataset were used. The distribution of log-of-the-odds (LOD) values are shown in Figure 5.2-6 for full-sibling pairs. The expected distributions overlap between full-sibling and unrelated individuals, which means that choosing a FPR that provides reasonable assurance no unrelated pairs will be falsely called full-siblings will result in an undesirably high FNR. For Skagit River Bull Trout, a LOD value = 8.0 (corresponding to FPR = 1.4 e^{-5}) results in a FNR = 0.15, meaning approximately 15 percent of true full-sibling comparisons would be misclassified as unrelated with an α=0.05 as the typical standard.

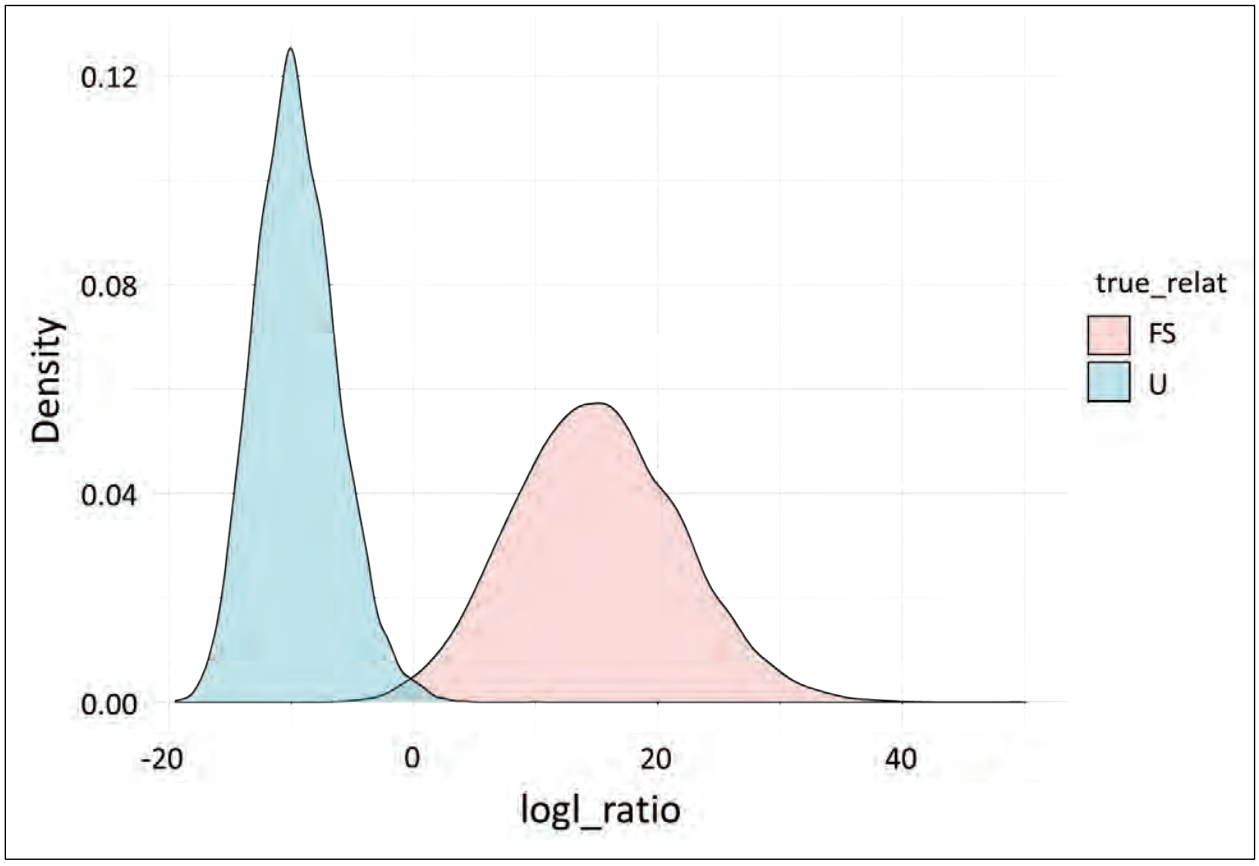


Figure 5.2-5. **Log likelihood ratios distribution for simulated true full siblings versus unrelated individuals based on Skagit River *S. confluentus* genotype Year 1 microsatellite data provided by WDFW and USFWS (Table 5.2-2). High overlap between full-siblings and unrelated fish suggests relatively low power to detect highly related individuals using the 16 microsatellites. Legend: FS=Full Sibling; U=Unrelated. High overlap between full-siblings and unrelated fish suggests relatively low statistical power to detect highly related individuals.**

Note that these estimated rates were based on all individuals analyzed (n=530), which would likely overestimate power for studies of "real-world" populations, because the sample used to estimate the allele frequencies was a mixed sample (i.e., across tributaries). A more realistic evaluation would consider collections from a single study area tributary, as opposed to considering potential comparisons between unrelated individuals across the entire Skagit River basin. The analysis was

therefore repeated using only collections from Big Beaver, Ruby, and Stetattle creeks in the study area. The FNR estimated for Big Beaver, Ruby, and Stetattle creek collections were 0.857, 0.868, and 0.95, respectively, meaning pedigree analysis is expected to result in more false relationship assignments than true assignments.

Understanding power to detect related individuals helped identify individual samples that might need to be removed from analysis to reduce violation of sampling assumptions. COLONY (Jones and Wang 2010) was used to screen collections for full sibling families, and based on power estimates above, applied probability of inclusion = 1.0 and a probability of exclusion = 0.99 to accept family classifications. Inclusion probability gives the probability that all individuals (in that family) are indeed full siblings from the same family. Exclusion probability is the probability those individuals are full siblings, and no other individuals are full siblings with this family. There is no accepted convention or criterion for identifying and removing related individuals from a dataset, although the criteria used here are more stringent than those referenced in literature pertaining to this Bull Trout dataset (Smith 2010). All full siblings but one were omitted from identified families within the collection (the presence of multiple representatives from the same family skews allele frequencies from true population proportions, creating a bias. Removing all but one sibling removes this bias).

5.2.2.3 Year 1 Population Determination

Similar to Rainbow Trout, PCA of allele frequencies (adegenet package) was used to examine genetic variation among collections. Data modeling suggested retention of approximately 15 PCs and 5 discriminant functions (k) would result in reliable partitioning of genetic variation among group clusters. Given an iterative DAPC data exploration, with the number of genetic group (k-means) clusters fixed at two (i.e., k=2), samples were partitioned into genetic groupings associated with Diablo/Gorge lakes and all other samples. With an additional cluster allowed (k=3), individuals were partitioned into (1) study area tributaries and some reservoir samples; (2) study area reservoir samples; and (3) samples from below Gorge Dam. With the allowance of fourth and fifth genetic clusters (k=4 and k=5), study area reservoir samples became split among the newly allowed clusters. No further refinement of study area samples was observed at higher numbers of clusters. A visualization of the k-means clustering at k=5 is shown on Figure 5.2-7. Clusters 3, 4 and 5 were predominantly individuals collected from Diablo and Gorge lakes (see Figure 5.2-8). Cluster 1 were study area tributary collections, constituting a majority of Ross Lake samples. Cluster 2 were individuals collected from below Gorge Dam.

As mentioned, collections submitted by WDFW were a part of evaluations intended to assess hybridization among Bull Trout, Dolly Varden, and Brook Trout. Reports pertaining to data noted that hybrids were observed within these collections (e.g., Small et al. 2013; Small et al. 2016). The study team was unable to directly ascribe clusters 3, 4, and 5 to hybridization among individuals or genetic introgression because (1) taxon-diagnostic alleles among taxa were unknown; (2) sample IDs for individuals WDFW considered hybrids were not provided; (3) the methods by which WDFW determined individuals to be hybrids was not provided; and (4) the selection strategy (if any) of field personnel collecting individuals “at large” from reservoirs was also not provided.

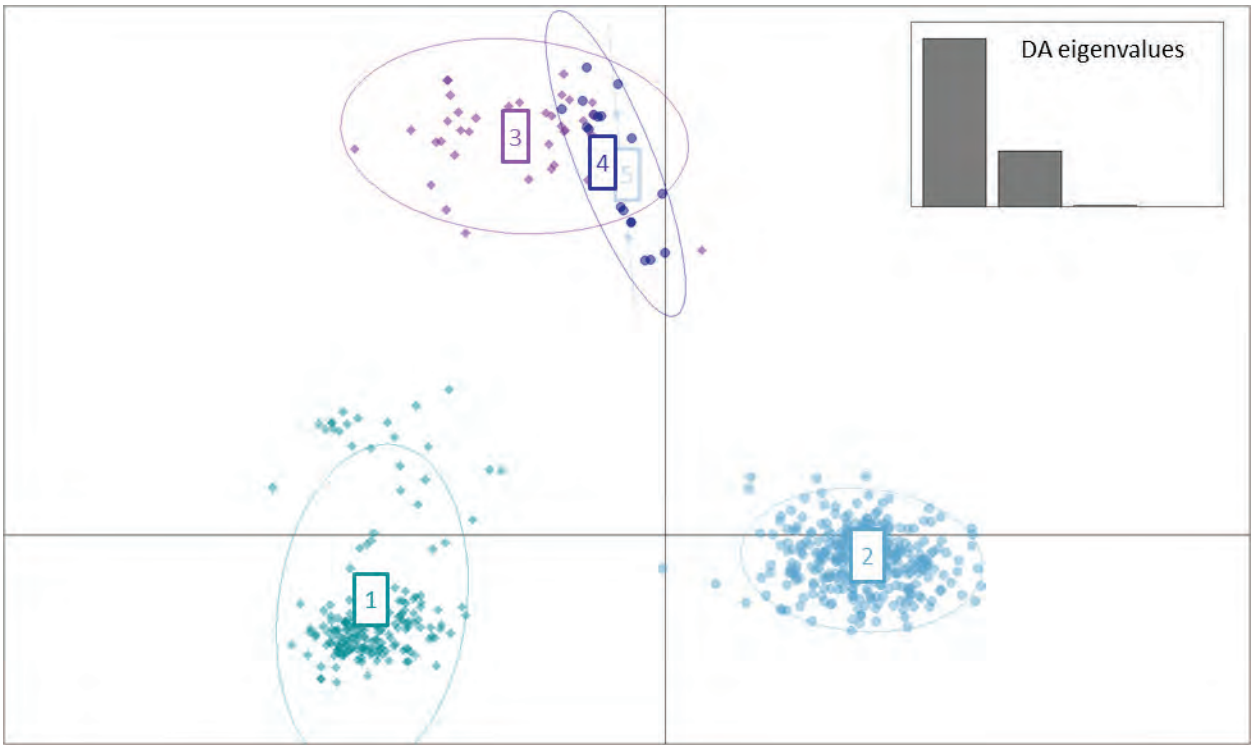


Figure 5.2-6. **Visualization of k-means clustering analysis at *k*=5 for Bull Trout individuals from previously reported microsatellite dataset at 1st and 2nd principal component axes. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in upper right corner shows first three eigenvalues. Cluster 1 were study area tributary collections and contained a majority of Ross Lake samples. Cluster 2 were individuals collected from below Gorge Dam. Clusters 3, 4 and 5 were predominantly individuals collected from Diablo and Gorge lakes (see Figure 5.2-8 for refined locations).**

Small sample sizes of *Salvelinus* spp. (median N=26) relative to *O. mykiss* (median N=45) highlighted limitations associated with balancing precision and bias. For instance, collections with fewer than 25 individuals are typically not recommended for analyses using microsatellite data, however, adopting this criterion for the *Salvelinus* spp. dataset would have resulted in exclusion of about 50 percent of Bull Trout collections from an already sparse dataset. The genetic groupings shown in Figure 5.2-8 also underscore the challenges associated with choosing which fish to retain in any given collection due to genetic admixture. All individuals in clusters 3, 4, and 5 were considered potentially admixed and omitted from the dataset prior to estimating genetic summary statistics for each collection. The resulting final dataset comprised n=530 samples (Table 5.2-2). The genotypes are saved in GENEPOP format and are available upon request.

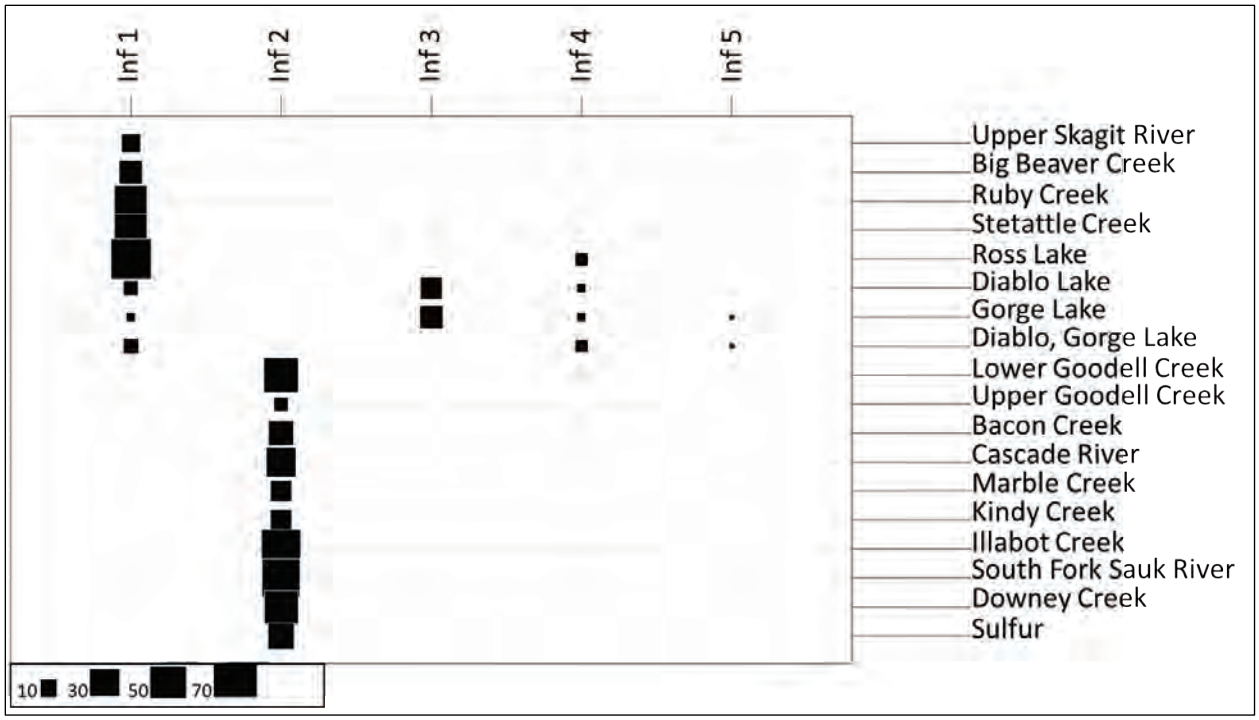


Figure 5.2-7. Genetic clusters visualized in Figure 5.2-7 aligned to each Bull Trout collection from previously reported microsatellite dataset. Size of boxes is scaled by sample count. Genetic clusters are organized by geographic location with upper Skagit collections at the top and lower Skagit at the bottom. Inferred clusters (inf 1-5) are the same as shown in Figure 5.2-7.

5.2.2.4 Year 1 Genetic summary statistics

Heterozygosity in the Bull Trout collections ranged from 0.337 to 0.467 within collections from study area tributaries (above Gorge Dam) and was 0.473 in the Ross Lake collection (Table 5.2-3. The collections from within the study area (above Gorge Dam) had lower heterozygosity than the collections from below Gorge Dam (Chi-square p-value = 0.0027). The attempt to reduce violation of HWE appeared successful, as mean F_{IS} across all collections was not statistically different from 0.00 (F_{IS}=0.008, 95 percent CI: -0.024-0.051). Each study area tributary collection (upper Skagit, Big Beaver, Ruby, Stetattle) did not deviate significantly from expectations. The Ross Lake collection was not in HWE, along with potentially several collections from below Gorge Dam, particularly Bacon Creek and Illabot Creek. Potential cause(s) of observed HWE deviations (e.g., data quality, inbreeding, population mixing) were not determined. We measured LD using log-likelihood (*G*) tests for all pairwise locus comparisons. Of the 1,680 comparisons (overall collections), 271 were significant at the α=0.05 level. No study area tributary collections (above Gorge Dam) had statistically significant LD tests using the adjusted table wide significance level α=0.0003. The Ross Lake collection had 11 significant LD tests out of 120. The greatest number of significant log-likelihood tests was observed for the Illabot Creek collection (16).

The estimated proportion of genetic variance explained by population structure (F_{ST}) across all Bull Trout collections was 0.188, and 0.03 among study area tributary collections, only. Pairwise log-likelihood (*G*) tests for population differentiation were not statistically significant between the

upper Skagit River, Big Beaver Creek, and Ruby Creek collections (adjusted nominal level 5 percent). The upper Skagit River collection was not differentiated from the Ross Lake collection, but the Ross Lake collection was differentiated from both the Big Beaver Creek and Ruby Creek collections. The Stetattle Creek collection was differentiated from all other study area collections. Note that the Marble Creek collection was not differentiated from any collection in the dataset except the South Fork Sauk River. This seemed anomalous, so results that follow exclude consideration of Marble Creek collection. All study area collections (above Gorge Dam) were differentiated from below Gorge Dam collections. Recall, F_{ST} is the proportion of genetic variation that is attributable to population subdivision with F_{ST}=0.00 reflecting no differences and F_{ST}=1.00 reflecting complete differentiation (i.e., all genetic diversity is partitioned among subpopulations). The F_{ST} estimated (pairwise) between the study area collections are shown in Table 5.2-4. For context, F_{ST} estimated from comparisons between the study area collections with those from below Gorge Dam ranged from a low of 0.207 to a high of 0.397, a result consistent with Smith (2010) (data not shown).

Table 5.2-3. Year 1 summary statistics for samples collected from Bull Trout in the Skagit River basin.

Collection	Sample Size (*n*)	F_{IS}[1]	H_s[2]	MNA[3]
Upper Skagit River	14	0.080	0.467	5.00
Big Beaver Creek	21	0.042	0.410	4.44
Ruby Creek	41	-0.021	0.384	4.75
Ross Lake	62	0.105	0.473	7.16
Stetattle Creek	41	-0.078	0.337	2.94
Goodell Creek	54	0.046	0.647	6.97
Bacon Creek	24	0.038	0.678	7.56
Illabot Creek	60	-0.050	0.634	7.44
Cascade River	33	0.033	0.662	8.19
Marble Creek	18	-0.080	0.679	6.94
Kindy Creek	17	0.016	0.689	7.19
S.F. Sauk River	54	-0.032	0.656	8.31
Downey Creek	44	0.010	0.709	9.88
Sulfur	27	0.035	0.607	6.13

1 $F_{IS:}$ estimated deviation from Hardy-Weinberg proportions.
2 $H_{S:}$ estimated expected heterozygosity within sub-populations (i.e., gene diversity).
3 MNA: is the mean number of alleles observed over all loci.

Table 5.2-4. Table of pairwise estimates of F_{ST} between the Project area collections of Bull Trout.

	Upper Skagit River	Big Beaver Creek	Ruby Creek	Ross Lake
Big Beaver Creek	0.001			
Ruby Creek	0.028	0.014		
Ross Lake	0.023	0.043	0.061	
Stetattle Creek	0.068	0.030	0.034	0.105

5.3 Year 2 Results

5.3.1 Year 2 Rainbow Trout (*Oncorhynchus mykiss*)

5.3.1.1 Year 2 Hybridization

Genotypes for 1,425 *Oncorhynchus* individuals from the study area were assessed for hybridization status. Eighty-seven percent of individuals collected possessed *Oncorhynchus* alleles at the three diagnostic loci present in the GTSeq panel (Figure 5.3-1). Conditional on resolution provided by diagnostic loci, non-hybrid *O. mykiss* represented 87 percent of collected samples. Of the 13 percent hybridized individuals observed (N=190), the hybrid index ranged from 1-6 (Figure 5.3-1). The geographic distribution of hybridized individuals observed is shown in (Figure 5.3-2). The hybridized *O. mykiss* were omitted from the dataset used for genetic analysis.

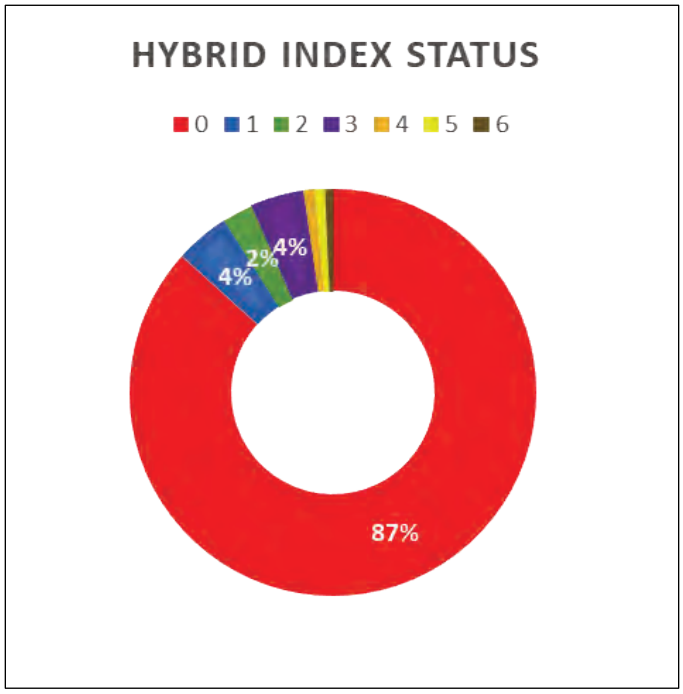


Figure 5.3-1. Proportional distribution of *O. mykiss* hybridization index. Non-hybrid *O. mykiss* make up 87 percent of collected samples, while the remaining 13 percent of hybridized individuals observed, ranged from 1-6 on the hybrid index. Proportions are conditional on resolution provided by diagnostic SNP loci present in SNP panel.

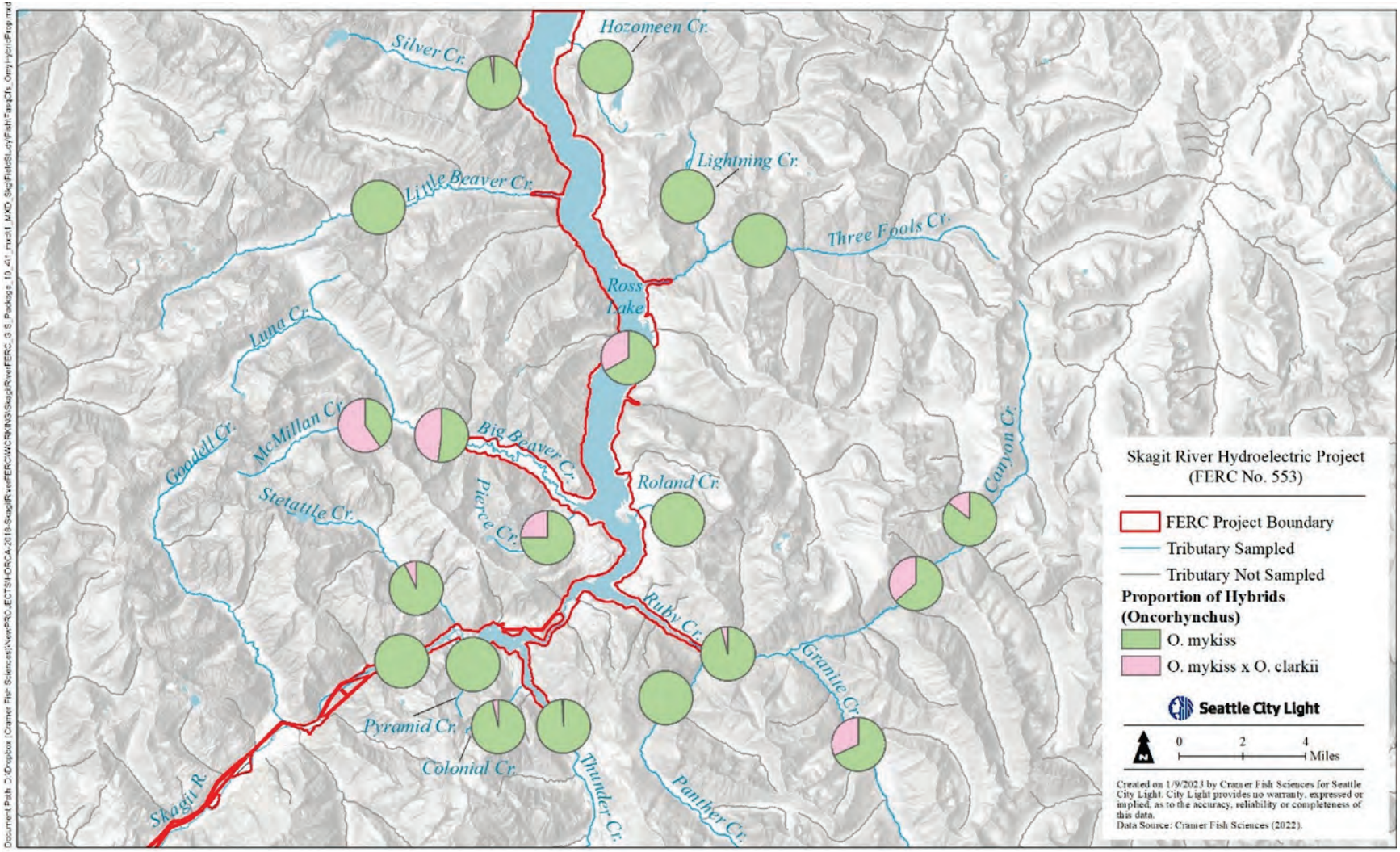


Figure 5.3-2. Year 2 map showing the distribution of the proportion of hybridized *Oncorhynchus* individuals collected across the Project area.

5.3.1.2 Year 2 Within and Among Population Diversity

After removal of 190 hybridized individuals and one individual for not having a verified collection location, 1,234 *O. mykiss* were retained for analysis (Table 5.3-1). The dataset was screened for neutral genetic loci that were uninformative (i.e., a minor allele frequency of less than or equal to 0.01). Seven loci matched this criterion and were removed from dataset (Omy104569114, OmyG3PD2371, Omyb9164, Omycarban1264, Omycyp17153, Omygadd45332, and Omysys1188), resulting in 235 neutral loci (per individual) retained for genotypes.

Table 5.3-1. Non-hybrid *O. mykiss* samples used for Year 2 genetic analysis.

No.	River/Stream Name	Collection Years[2]	Sample Size (*n*)
1	Silver Creek	2022	44
2	Hozomeen Creek	2022	3
3	Little Beaver Creek	2019, 2021, 2022	73
4	Lightning Creek	2018, 2020, 2022	143
5	Three Fools Creek	2019, 2020, 2021, 2022	103
6	Big Beaver Creek	2019, 2022	52
7	McMillan Creek	2021, 2022	6
8	Pierce Creek	2022	3
9	Roland Creek	2022	8
10	Ross Lake	2020, 2022	4
11	Ruby Creek	2019, 2022	185
12	Canyon Creek	2018, 2019, 2022	94
13	North Fork Canyon Creek	2022	6
14	Panther Creek	2022	35
15	Granite Creek	2019, 2022	109
16	Colonial Creek	2022	53
17	Thunder Creek	2019, 2022	98
18	Stetattle Creek	2021, 2022	125
19	Pyramid Creek	2022	60
20	Gorge Lake[1]	2019	30

1 WDFW genotype data
2 CFS collections from 2022. All other dates shown are USGS collected samples

Prior to conducting genetic analysis on populations, the populations to analyze must be determined. A heuristic assessment of coherent genetic groups was conducted using DAPC. An initial exploratory DAPC used 200 genetic PCs and number of clusters (k) from 1-15, which considered 98.3 percent of observed genetic variance in the dataset. The initial DAPC was evaluated further, as importantly, retention of large numbers of PCs with respect to the number of individuals analyzed can over-fit the discriminant functions. If this occurs, individual membership in selected k clusters can become statistically unreliable, as discriminant functions could become flexible enough to discriminate any number of clusters, overinflating best-fitting clusters. The

trade-off between power of discrimination and overfitting can be measured by the a-score (see Section 4.4.4.2).

Implementation of the a-score procedure repeated DAPC on the dataset using from 1 up to 50 PCs sequentially, with seven PCs estimated to optimize the proportion of successful reassignment corrected for the number of retained PCs (data not shown). DAPC was then rerun on the *O. mykiss* dataset using seven retained PCs (instead of the initial 200), which considered 24.4 percent of the observed genetic variance in the dataset. When considering BIC-based selection of various possible number of clusters (k), the primary infliction point was for k=4 (i.e., four genetic clusters Figure 5.3-3), which provided a data driven starting point for the potential number of populations present in the study area. At the risk of causing confusion, there were a series of preliminary genetic analyses conducted on iterations of clustering *O. mykiss* individuals that will not be detailed here. Instead, the logic and reasoning will be described in brief on how both k and classification of individuals to populations was achieved using the observed data.

While individual probabilities for cluster membership were statistically reliable at k=4, genetic cluster 1 was inconsistent with this cluster representing a single population given subsequent genetic analysis (e.g., HWE). At k=5, that same genetic cluster 1 split into two genetic clusters, labeled as cluster 3 and 5 in Figure 5.3-4 (note that cluster number labels are arbitrary and cannot be specified in DAPC). Visualization of the membership probabilities for all fish analyzed showed that individuals from clusters 1, 2, and 4 were distinctive, and individuals were attributed to cluster 3 and 5 with varying probability (Figure 5.3-5). Note that individual fish are displayed approximate north to south, with Silver Creek starting on the right side of Figure 5.3-5 and Stetattle Creek ending on the left. Itemization of where each individual *O. mykiss* resides with respect to membership probability values are not shown. Cluster 1 is Little Beaver Creek, cluster 2 is Pyramid Creek, cluster 4 is Three Fools Creek, with all remaining Project *O. mykiss* residing in either clusters 3 or 5 (Table 5.3-2). At k=6, membership probabilities did not improve classification of individuals (data not shown), so k=5 was determined to be the logical categorization based on discriminant analysis of genotypes. Following data exploration using DAPC and completion of population analysis, output from a STRUCTURE analysis was qualitatively compared to DAPC results. Given consideration of population number (K) from 2 to 10, and three technical replicates at each K, the Evanno method suggested that K=6 was a best fit for the dataset (data not shown). Cutthroat Trout controls and *O. mykiss* hybrids received their own population (K), with some undetected hybridized *O. mykiss* individuals (varying levels) observed in the dataset. Therefore, the findings from DAPC were corroborated by STRUCTURE, with both methods showing K=5 as a best representation of underlying genetic variation. STRUCTURE placed Little Beaver Creek, Three Fools Creek, and Pyramid Creek in their own K, with two other widely distributed K's associated with the aforementioned DAPC clusters 3 and 5.

At k=5, genetic clusters 3 and 5 were still inconsistent with these clusters representing single populations (data not shown). Therefore, the genetic dataset was partitioned by both DAPC cluster and geographic location for subsequent population analysis to maximize consideration of genetic variation observed. HWE was estimated for each genetic locus (235 loci) within each population (28 populations). In this data configuration, a majority of loci conformed to HWE expectations, and no locus failed HWE across all populations (Figure 5.3-6), indicating the genetic loci were suitable for population analysis of study area *O. mykiss*. Relatedness among individuals within each population was estimated using Rxy metric. Mean Rxy was negative for all populations

except Lightning and Three Fools Creeks, with confidence intervals overlapping zero (Figure 5.3-7). The small number of individuals observed with Rxy greater than 0.5 were not omitted from data analysis. Genetic diversity (both observed heterozygosity and expected heterozygosity under Hardy-Weinberg model) is shown in Figure 5.3-8 for populations with greater than 15 samples. Diversity is highest in Stetattle Creek and lowest in Three Fools Creek, with expected heterozygosity higher than observed heterozygosity. This distribution of diversity resulted in positive F_{IS} values for many populations (Table 5.3-2), meaning there was a reduction in heterozygosity observed from what was expected under assumptions of Hardy-Weinberg model. Another measure of genetic diversity, allelic richness, was highest in Stetattle Creek (1.321) and lowest in Three Fools Creek (1.116).

Table 5.3-2. Non-hybrid *O. mykiss* samples used for genetic analysis. Bold F_{IS} values were statistically significant from zero.

No.	Location/DAPC cluster[1]	n	F_{IS}	Allelic Richness
1	Silver Creek-3	44	**0.058**	1.303
2	Hozomeen Creek-3	3	**-0.067**	NA
3	Little Beaver Creek-1	73	**0.062**	1.161
4	Lightning Creek-3	138	**0.04**	1.269
5	Three Fools Creek-4	108	0.036	1.116
6	Big Beaver Creek-3	46	0.021	1.291
7	Big Beaver Creek-5	6	-0.057	1.271
8	McMillan Creek-3	6	**0.104**	NA
9	Pierce Creek-3	3	**0.061**	NA
10	Roland Creek-3	8	0.008	NA
11	Ross Lake-3	4	-0.111	1.309
12	Ruby Creek-3	174	0.019	1.299
13	Ruby Creek-5	11	0.026	1.278
14	Canyon Creek-3	47	**0.032**	1.294
15	Canyon Creek-5	47	**0.034**	1.276
16	NF Canyon Creek-5	6	-0.023	NA
17	Panther Creek-3	3	-0.071	1.268
18	Panther Creek-5	32	0.016	1.222
19	Granite Creek-3	41	**0.022**	1.287
20	Granite Creek-5	68	**0.052**	1.266
21	Colonial Creek-3	53	0.015	1.299
22	Thunder Creek-3	96	**0.023**	1.303
23	Thunder Creek-5	2	**-0.14**	NA
24	Stetattle Creek-3	59	0.022	1.306
25	Stetattle Creek-5	66	**0.049**	1.321
26	Pyramid Creek-2	60	**0.048**	1.252
27	Gorge Lake-3 [2]	29	0.01	NA
28	Gorge Lake-5 [2]	1	NA	NA

1 Populations are label by location description and DAPC cluster membership.
2 WDFW genotype data

Pairwise estimates of F_{ST} were calculated amongst populations with greater than 15 samples (Figure 5.3-9). These measures can be interpreted as a genetic distance. Little Beaver, Three Fools, and Pyramid creeks were the most divergent Project populations, corroborating the DAPC analysis; however, all pairwise estimates of F_{ST} except one (Granite-1 versus Canyon-1) were statistically significant (i.e., non-zero). There were some opportunities to compare populations within the same connected tributary. Comparisons were possible for 1) Lightning and Three Fools creeks and 2) Ruby, Canyon, Granite, and Panther creeks. As mentioned, Three Fools Creek was distinctive, so was divergent from the downstream population in Lightning Creek. Ruby, Canyon, and Granite populations from genetic cluster 3 had the lowest F_{ST} values observed. Panther Creek genetic cluster 5 was divergent from other populations in this tributary. Canyon Creek genetic cluster 5 was also more divergent from Ruby, Canyon, and Granite genetic cluster 3 then this population was from Granite Creek cluster 5. Additionally, Ruby, Canyon, and Granite genetic cluster 3 was more similar to adjacent tributaries (e.g., Big Beaver) than to genetic cluster 5 within the same tributary. The global underlying distance pattern observed amongst comparisons between genetic clusters 3 and 5 was that F_{ST} were smaller between cluster 3 populations, irrespective of location, than between cluster 3 and cluster 5 populations. In contrast, while Canyon-5 and Granite-5 exhibited a small F_{ST}, comparisons between cluster 5 populations (Panther-5, Canyon-5 and Granite-5, Stetattle-5) tended to be large.

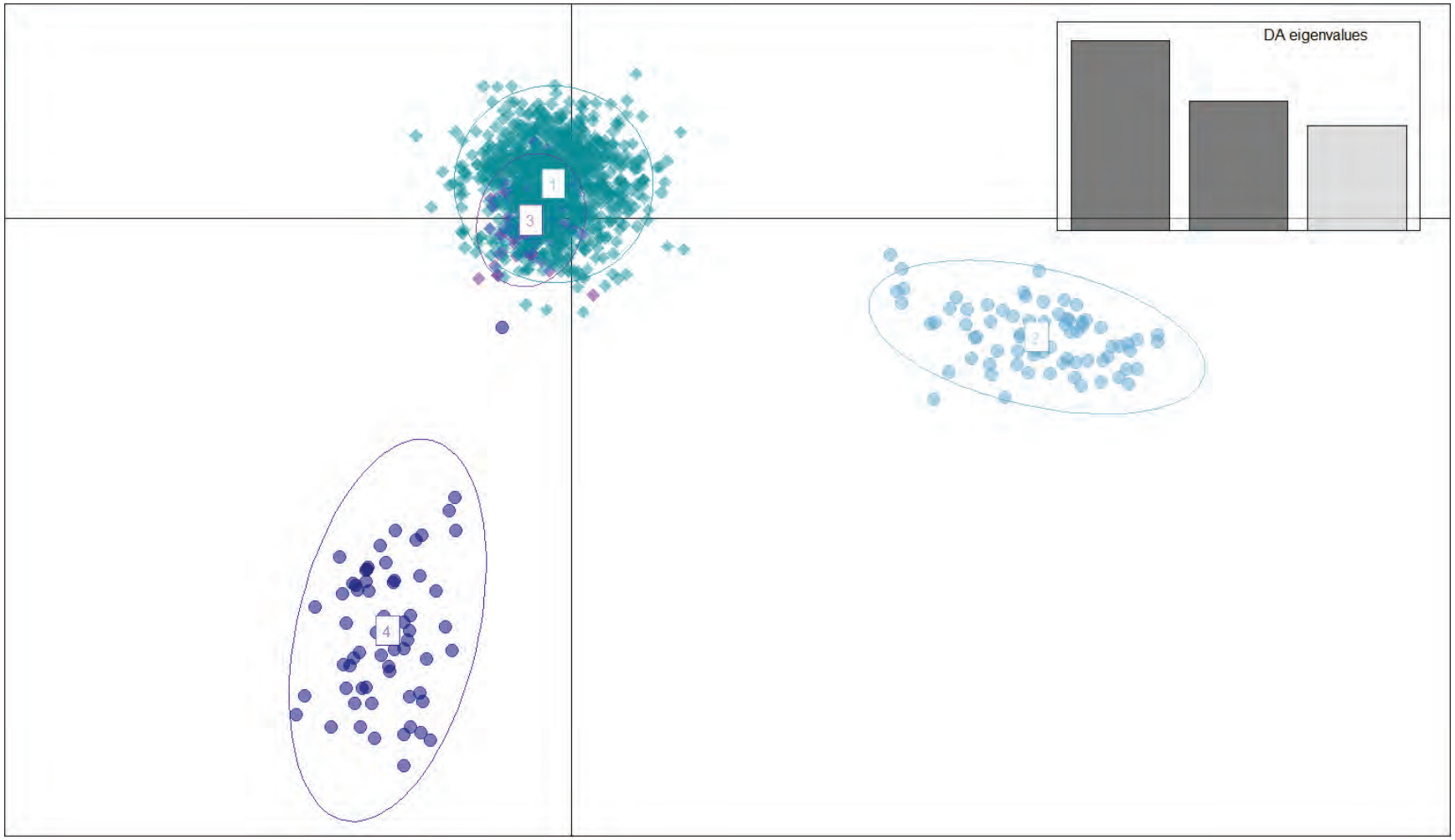


Figure 5.3-3. **Visual representation of DAPC analysis on 1st and 2nd principal components axes of Study Area *O. mykiss* for *k*=4 genetic clusters. Cluster 1 contained a majority of Study Area *O. mykiss*, cluster 2 was Little Beaver Creek, cluster 3 was mostly Three Fools Creek, and cluster 4 was Pyramid Creek.**

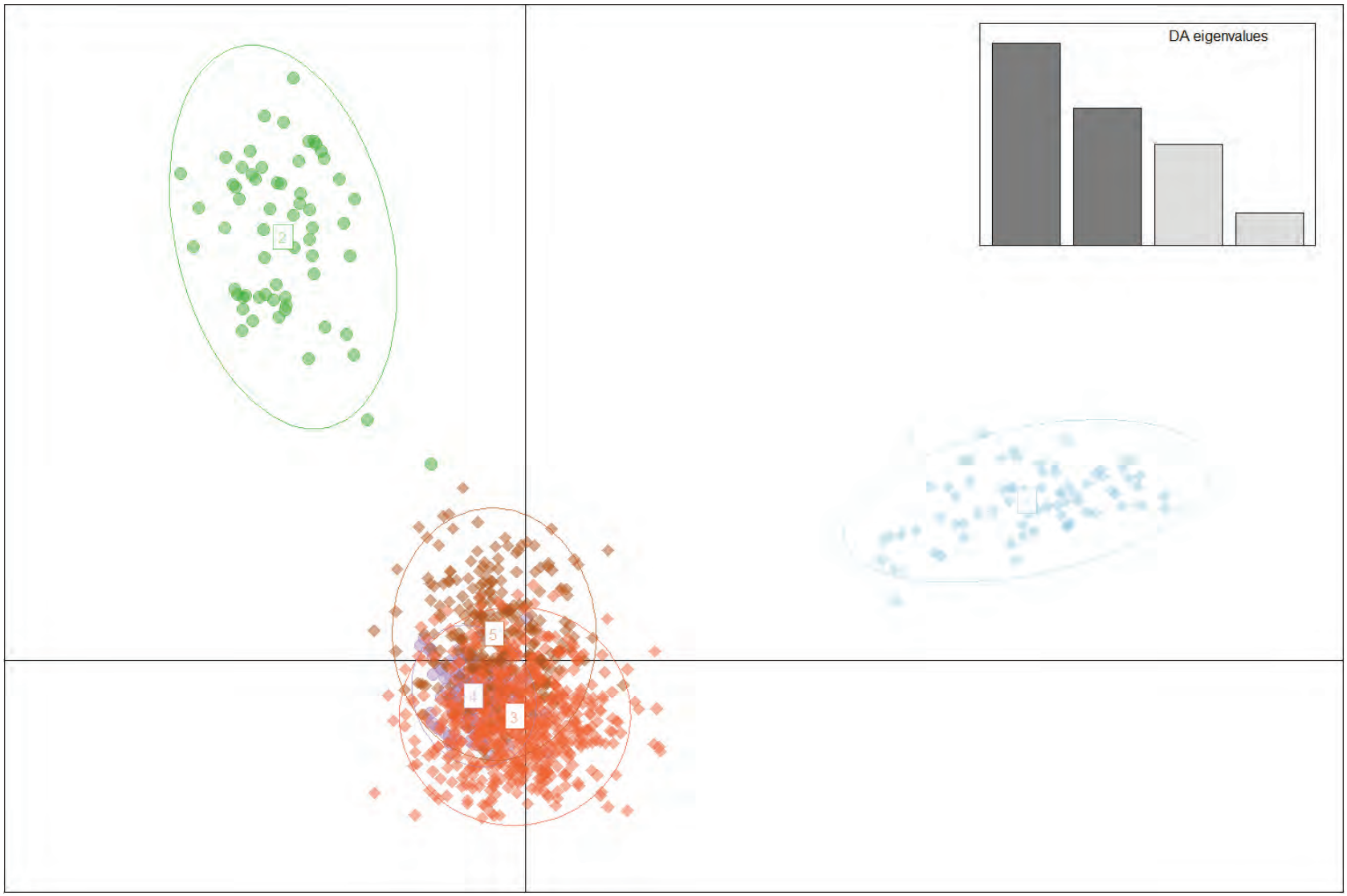


Figure 5.3-4. **Visual representation of DAPC analysis on 1st and 2nd principal components axes of Project *O. mykiss* for *k*=5 genetic clusters. Cluster 1 was identified as Little Beaver Creek, cluster 2 as Pyramid Creek, cluster 4 as majority Three Fools Creek. Individuals assigning to Clusters 3 and 5 were widely distributed across remaining tributaries in Study Area.**

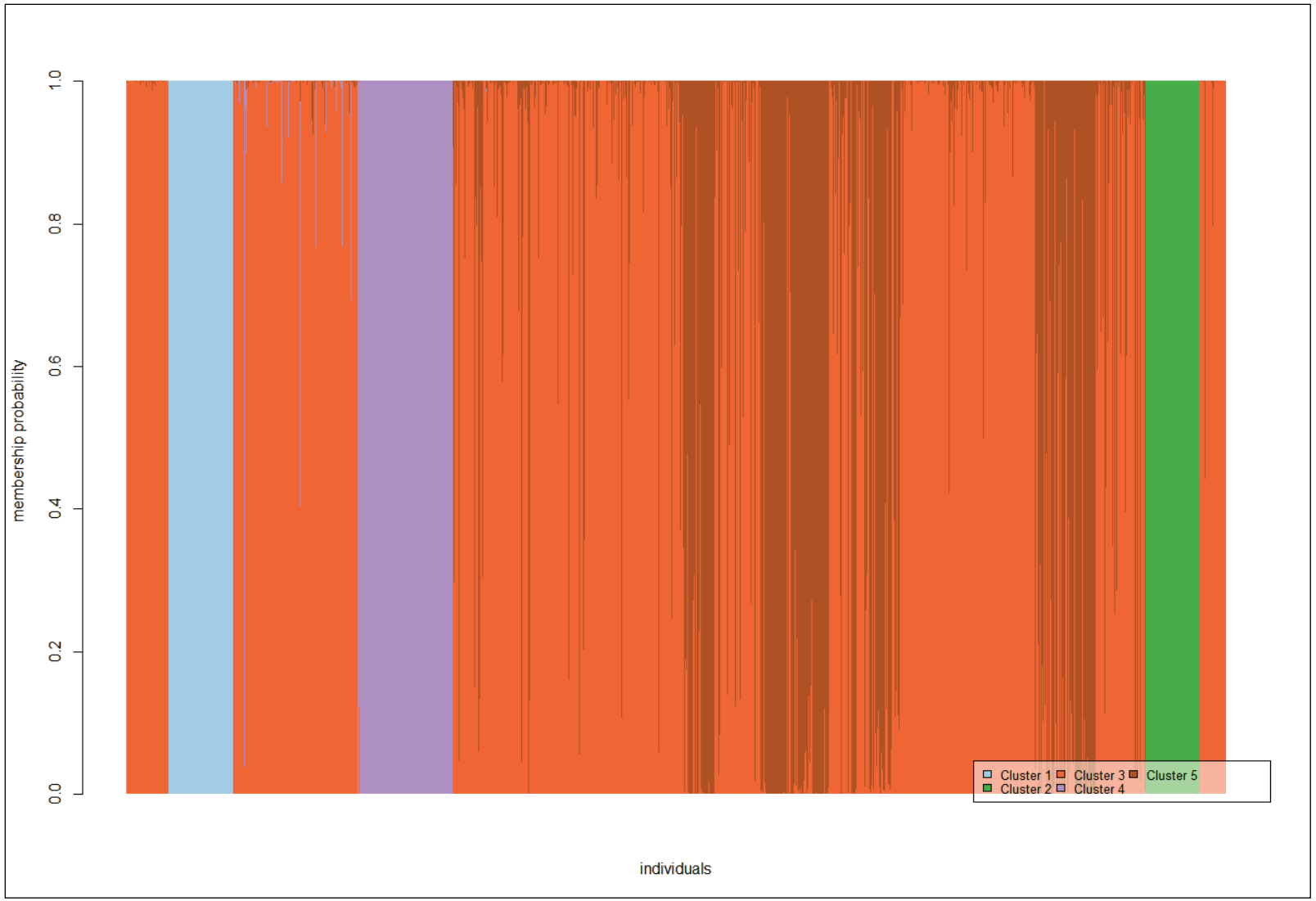


Figure 5.3-5. **Visual representation of membership probabilities given the same *O. mykiss* k=5 genetic clusters shown in Figure 5.3-4. Individual fish are displayed approximate north to south (Silver Creek starting on the right and ending with Gorge Lake on left). Cluster 1 was Little Beaver Creek, cluster 2 was Pyramid Creek, and cluster 4 was majority Three Fools Creek.**

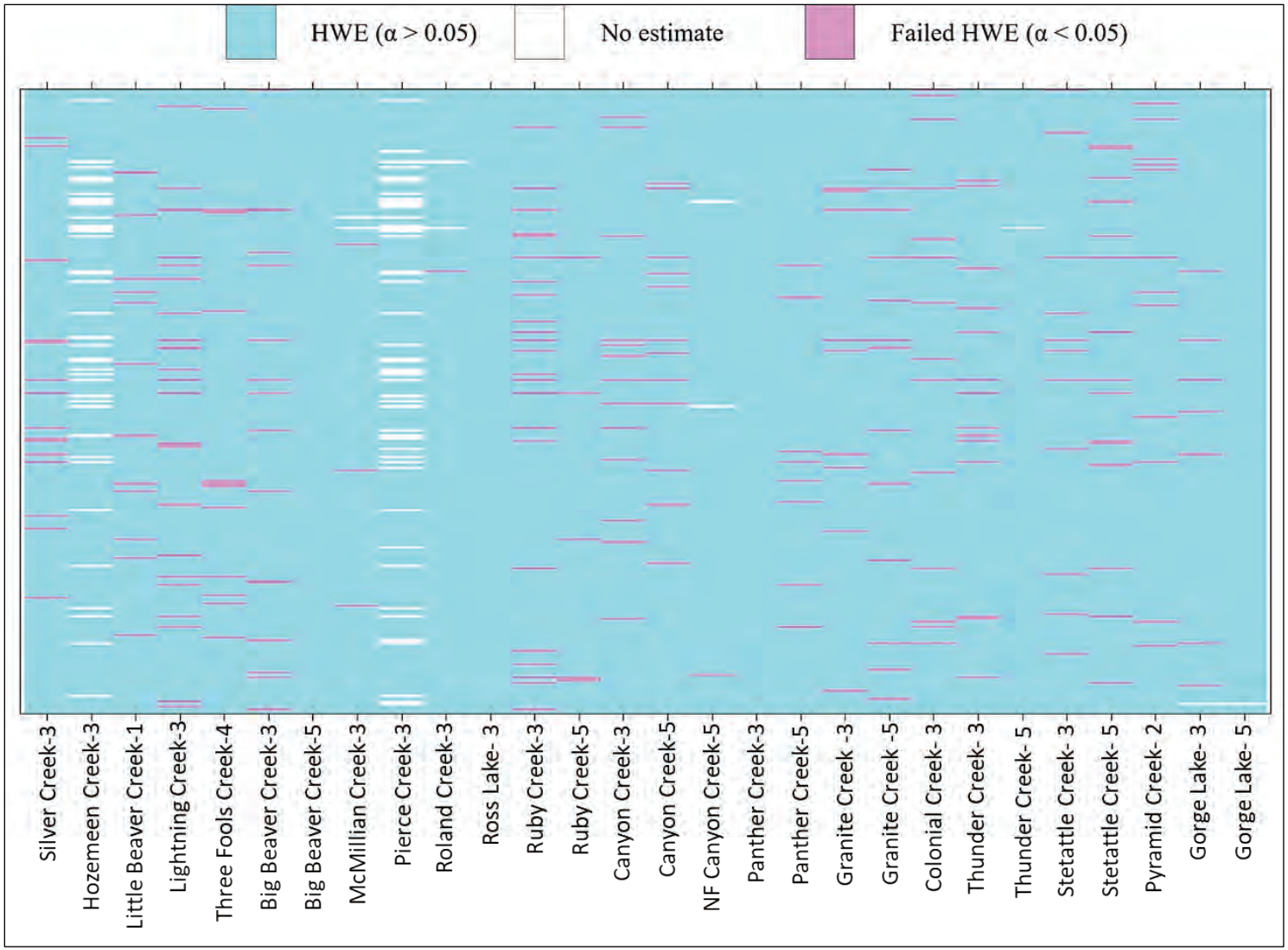


Figure 5.3-6. Visual representation of HWE estimates, with locus by population combination failing test at α less than 0.05 shown in pink. In this data configuration, a majority of loci conformed to HWE expectations, and no locus failed HWE across all population.

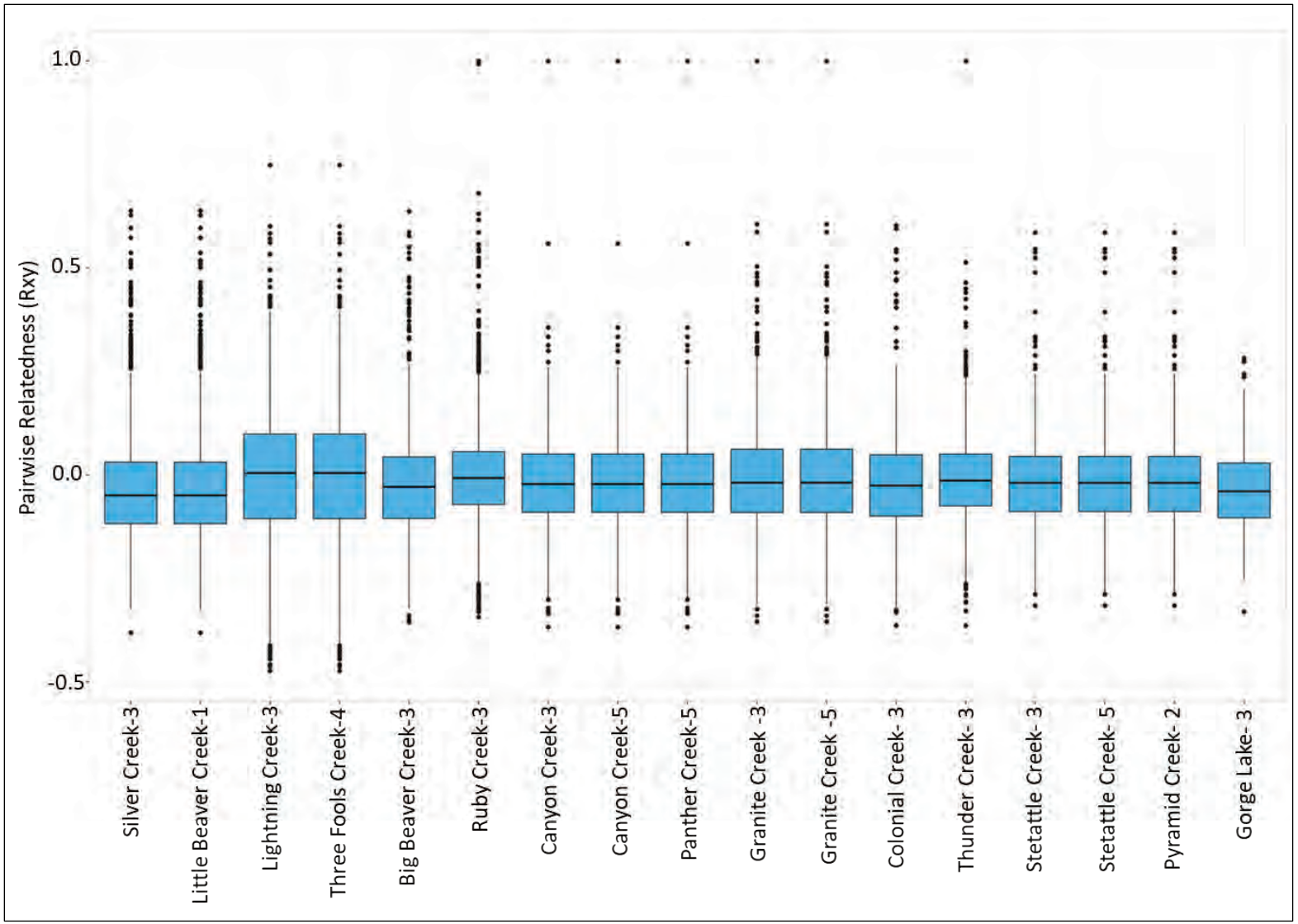


Figure 5.3-7. Pairwise relatedness (Rxy) between individual *O. mykiss* within each Project population containing greater than 15 samples. For all populations except Lightning and Three Fools Creeks, the mean Rxy was below zero. An Rxy = 0.0, 0.25, and 0.50 equates to individuals being unrelated, half siblings, and full siblings, respectively.

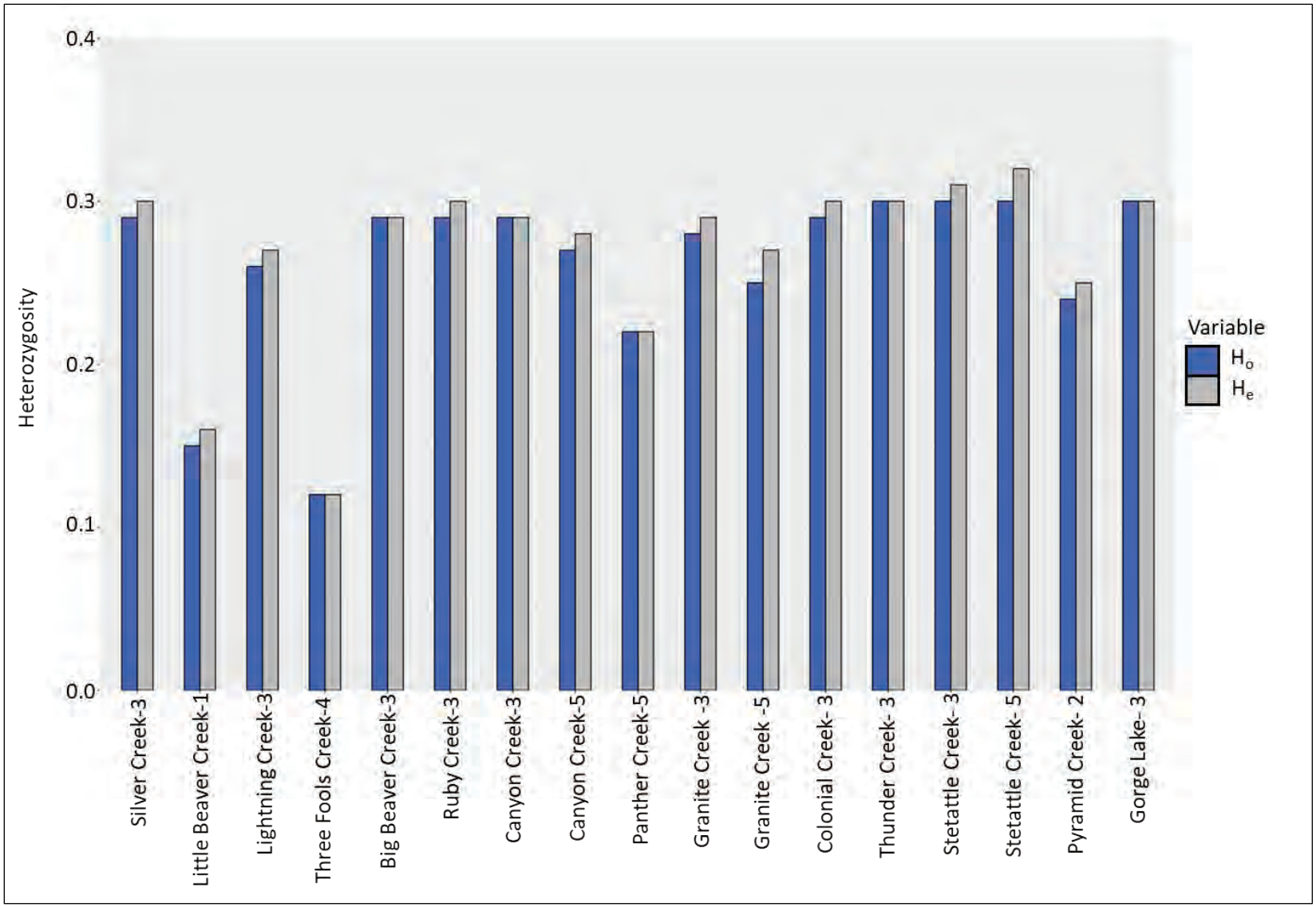


Figure 5.3-8. Genetic diversity (mean observed and mean expected heterozygosity) for Project *O. mykiss* populations containing greater than 15 samples.

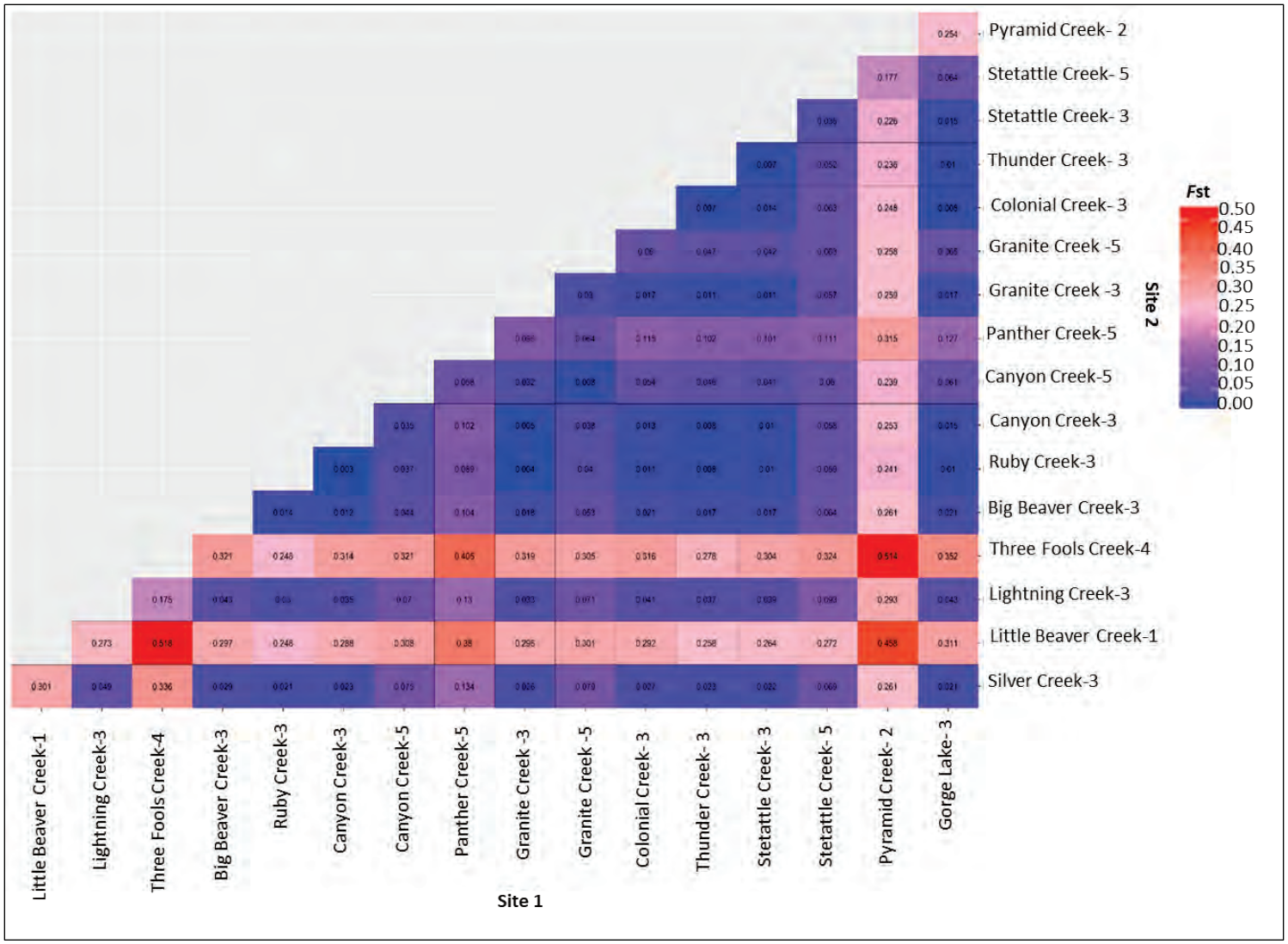


Figure 5.3-9. **Pairwise estimates of F_{ST} for Study Area *O. mykiss* populations containing greater than 15 samples. All pairwise F_{ST} estimates were statistically significant except for Granite-3 versus Canyon-3. F_{ST} values shown are the actual F_{ST} quantities, and not the significance of each pairwise test.**

5.3.1.3 Year 2 Scale Age Determination

During the 2022 field season, scales were collected from 407 *Oncorhynchus* individuals, of which 84 were aged. The 84 *Oncorhynchus* samples contained six age classes (0+, 1+, 2+, 3+, 4+, and 5+) with most individuals consisting of younger (age 2 or younger) age classes (Table 5.3-3). There was considerable overlap of fork lengths between age class 0+ and 1+ and 1+ and 2+ (Figure 5.3-10). Less fork length overlap was observed in older fish, but sample size was smaller.

Table 5.3-3. Size and age summaries for scale-determined ages of *Oncorhynchus* individuals.

Age Class	*n*	Min	Max	Mean	Std Dev
0+	20	55	95	69	13.9
1+	32	65	165	110.9	26.4
2+	20	105	275	168	37.9
3+	4	165	305	237.5	60.8
4+	7	305	375	349.3	21.5
5+	1	335	335	335	N/A

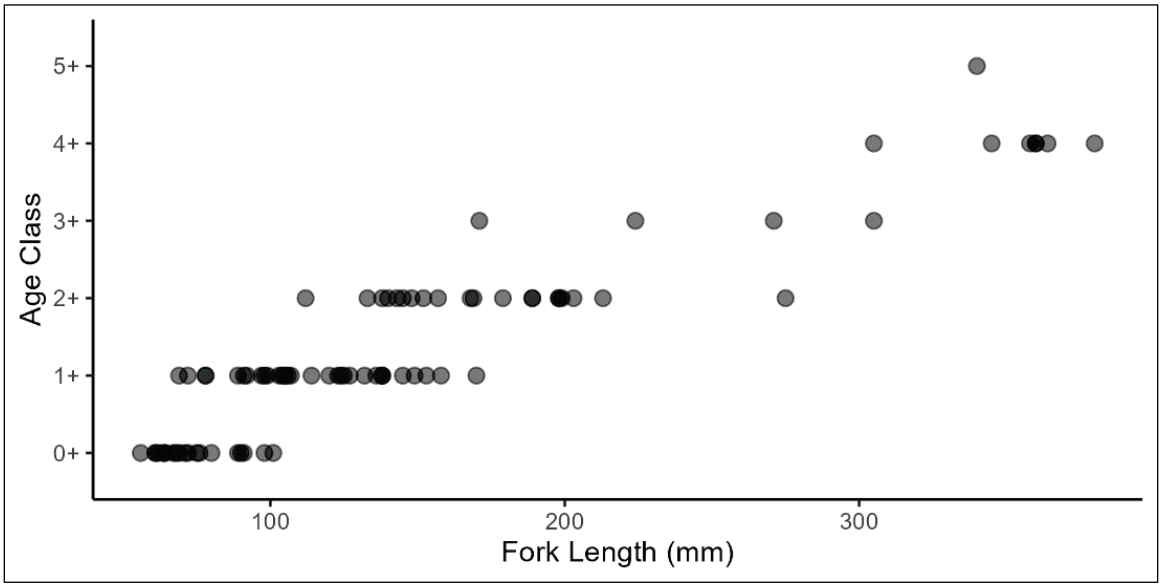


Figure 5.3-10. Summary of age, based on scale analysis, and fork length (mm) of 84 *Oncorhynchus* captured in 2022.

5.3.1.4 Length-at-Age Key Assignment

A total of 484 un-aged *Oncorhynchus* samples were assigned ages based on the length-at-age key for a total of 508 individuals (Table 5.3-4). There was considerable overlap of fork lengths between adjacent ages for classes 0+, 1+, and 2+; assignment of larger fish was difficult due to few aged fish above 200 mm. (Figure 5.3-11).

Table 5.3-4. Size and age summaries for aged and age-assigned *Oncorhynchus* individuals.

Age Class	*n*	Min	Max	Mean	SD
0+	85	55	103	80.06	14.23
1+	248	67	174	117.6	23.43
2+	134	105	295	171.8	41.73
3+	27	168	313	236.37	41.17
4+	12	305	380	339.75	26.24
5+	2	340	342	341	1.41

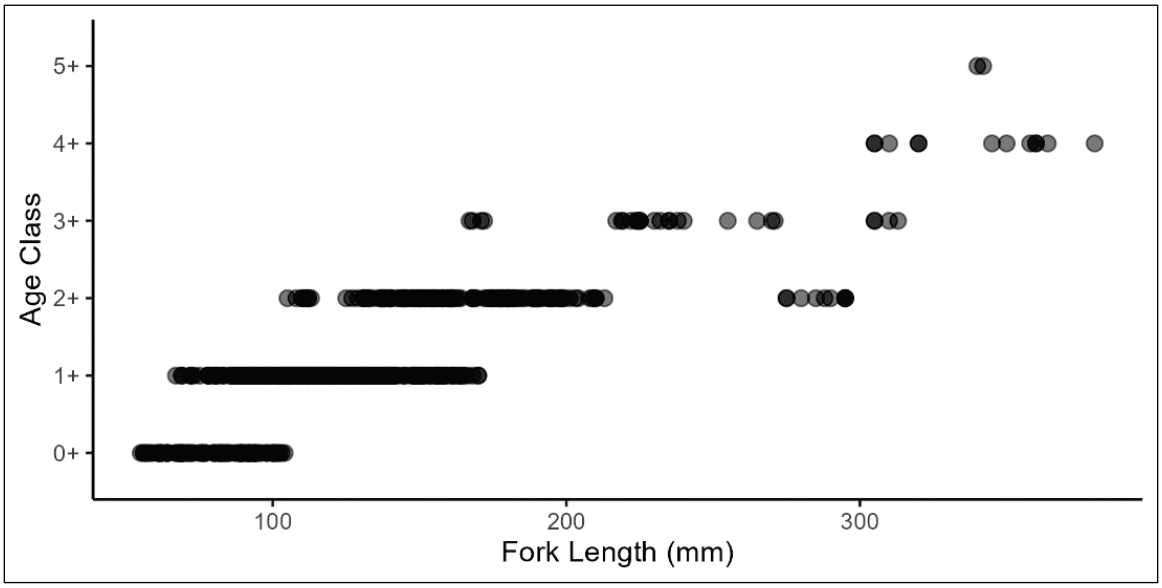


Figure 5.3-11. Summary of age, based on scale analysis and age-assignment, and fork length (mm) of 508 *Oncorhynchus*.

5.3.1.5 Year 2 Effective Population Size

There are various ways that population by age data can be parsed for estimation of effective population size (N_b). Given the population analysis above, genetic clusters 3 and 5 would be analyzed separately for each single age cohort. However, the quantity of samples collected and aged from 2022 were insufficient to achieve this configuration. Alternatively, genetic cluster 3, excluding Lightning Creek, had pairwise F_{ST} of approximately 0.02, which is a theoretical threshold for genetic drift connectivity. Therefore, for the initial calculations estimating of annual values for N_b of Project *O. mykiss*, age-1 individuals from genetic cluster 3, excluding Lightning Creek, were combined into a single sample (N=110). Additionally, N=15, N=34, and N=33, age-1 individuals from Little Beaver Creek, Pyramid Creek, and Stetattle Creek (genetic cluster 5) were analyzed as separate populations.

The annual effective population size (N_b) from the amalgamated genetic cluster 3 was 394.9 (95 percent CI 321.0-508.5). Little Beaver Creek (genetic cluster 1) age-1 cohort had an estimated N_b of 106.2 (95 percent CI 53.8-1144.1). Pyramid Creek (genetic cluster 2) age-1 cohort had an

estimated N_b of 13.5 (95 percent CI 12.5-14.7). Stetattle Creek (genetic cluster 5) age-1 cohort had an estimated N_b of 28.7 (95 percent CI 26.2-31.5).

5.3.1.6 Year 2 Haplotype Diversity

The two most frequent haplotypes observed at OMY5 loci were 1-4-3-3 and 3-1-1-1. The haplotypes Pearse et al. (2014) found associated with juvenile propensity to emigrate (exhibit anadromous behavior) were not observed in Project *O. mykiss (*i.e., 4-3-3-1 or 1-1-1-3). The most common OMY5 haplotype overall (haplotype 3) was also the most common haplotype observed in 22 of 28 populations in the study area (Figure 5.3-12).

Table 5.3-5. Diversity of OMY5 haplotypes for Project *O. mykiss* individuals.

Haplotype	OmyR14589	OmyR19198	OmyR24370	OmyR33562	Hap.freq
1	1	1	1	1	0.000
2	1	4	1	3	0.002
3	1	4	3	3	0.867
4	3	1	1	1	0.123
5	3	1	1	3	0.001
6	3	4	1	3	0.007
7	3	4	3	3	0.000

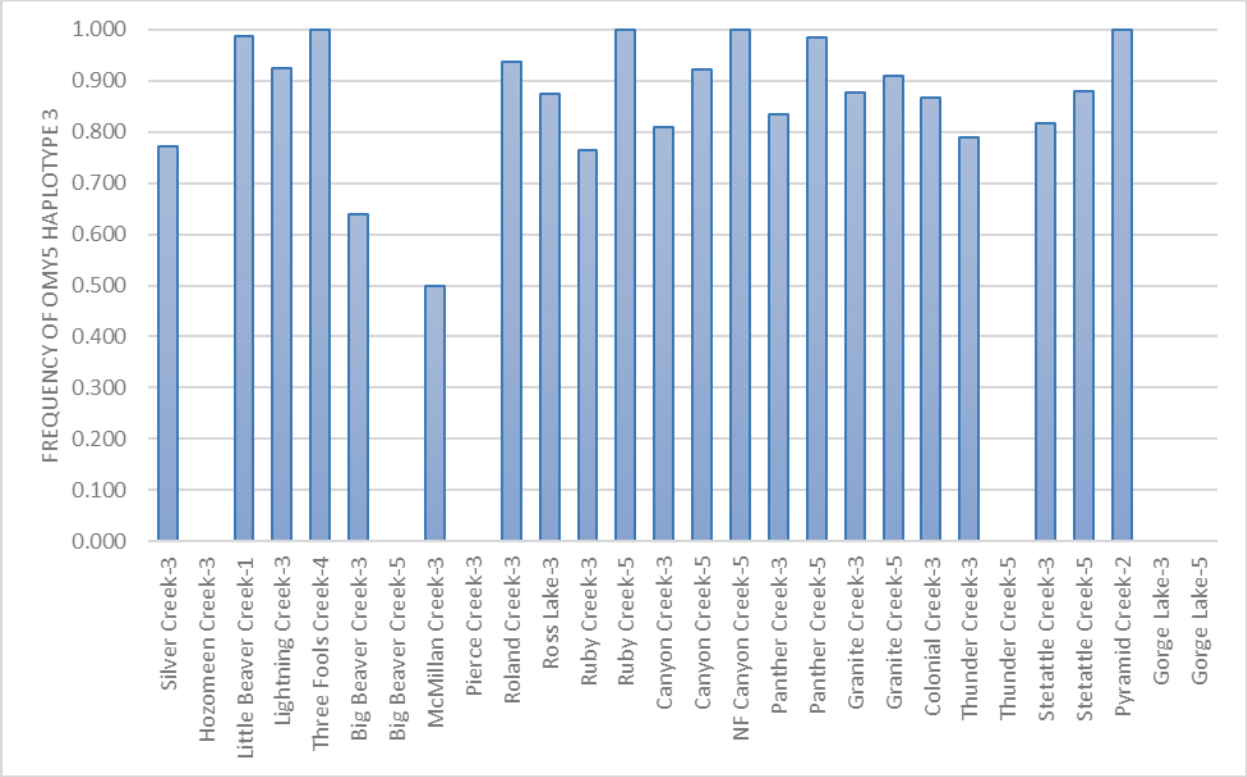


Figure 5.3-12. Frequency of the most common OMY5 haplotype (haplotype 3) in Project *O. mykiss* populations.

5.3.1.7 Year 2 Above and Below Project Population Analysis

For population analysis of above- and below-Project (Gorge Dam) *O. mykiss*, data provided by WDFW was analyzed along with data generated by CFS. Note that *O. mykiss* data provided by WDFW has been subject to numerous previous analyses, so the intent of CFS using the data was to provide context for diversity and distance (F_{ST}) values observed, in addition to the relative magnitude of genetic differentiation. CFS requested data from upper Skagit River, Goodell Creek, Marblemount Hatchery, lower Cascade River, and Finney Creek to incorporate into analysis representative population data for Skagit River *O. mykiss*. Importantly, the majority of WDFW data was not generated using the CRITFC-developed 354 SNPs GTSeq panel used by CFS for analysis. Rather, the WDFW data consisted of a (previous iteration) smaller 180 SNP locus panel. Therefore, CFS omitted data (nonoverlapping loci) from the total dataset for *O. mykiss* to form a complimentary set of data to use with smaller WDFW dataset. For the combined dataset, 178 loci were considered informative (minor allele frequency greater than 0.01).

Implementation of the DAPC a-score procedure on the "above-below" dataset estimated retention of 12 PCs optimized the proportion of successful reassignment corrected for the number of retained PCs (data not shown). DAPC was then rerun on the above-below *O. mykiss* dataset using 12 retained PCs, which considered 30.8 percent of the observed genetic variance in the dataset. When considering BIC based selection of various possible number of clusters (k), the primary inflection point was unclear, but k=5 (i.e., five genetic clusters; Figure 5.3-13) provided the highest k that both exemplified the underlying genetic principal components and resulted in reliable membership probabilities. Higher k did not change the general data pattern and merely subdivided populations on either side of above/below boundary (data not shown). Itemization of where each individual *O. mykiss* resides with respect to membership probability values are not shown here. Cluster 1 was predominantly Three Fools Creek with some Lightning Creek individuals included; most Project *O. mykiss* populations resided in cluster 2; cluster 3 was primarily Pyramid Creek; all but one individual from below Project populations resided in cluster 4, and cluster 5 was Little Beaver Creek. Rendered in 2-dimensions, the primary axis (x-axis) of Figure 5.3-13 pertains to above and below the Project. Above-Project populations (excluding Pyramid Creek) were to the left of the origin and Pyramid Creek, and below-Project populations were to the right of the origin. The second axis was driven by Little Beaver Creek genetic differentiation.

To summarize population diversity, the same location by genetic cluster designations used above in Section 5.3.1.1 were retained here, with the addition of upper Skagit River, Goodell Creek, Marblemount Hatchery, lower Cascade River, and Finney Creek populations. The upper Skagit River and Stetattle Creek collections had the highest diversity, with below-Project populations having observed heterozygosity greater than or equal to 0.3 (Figure 5.3-14). Above-Project populations had lower observed heterozygosity relative to below-Project populations. Note that the Marblemount Hatchery was the only population with greater observed heterozygosity than expected heterozygosity, suggesting these individuals were outbred relative to Hardy-Weinberg expectations. Pairwise estimates of F_{ST} were calculated amongst populations with greater than 15 samples. The observed magnitudes of F_{ST} were distributed as expected given DAPC (Figure 5.3-15). While accounting for the highly divergent populations (Little Beaver, Three Fools, and Pyramid creeks), F_{ST} were generally larger for comparisons between above and below populations than for comparisons amongst the populations from below Gorge Dam.

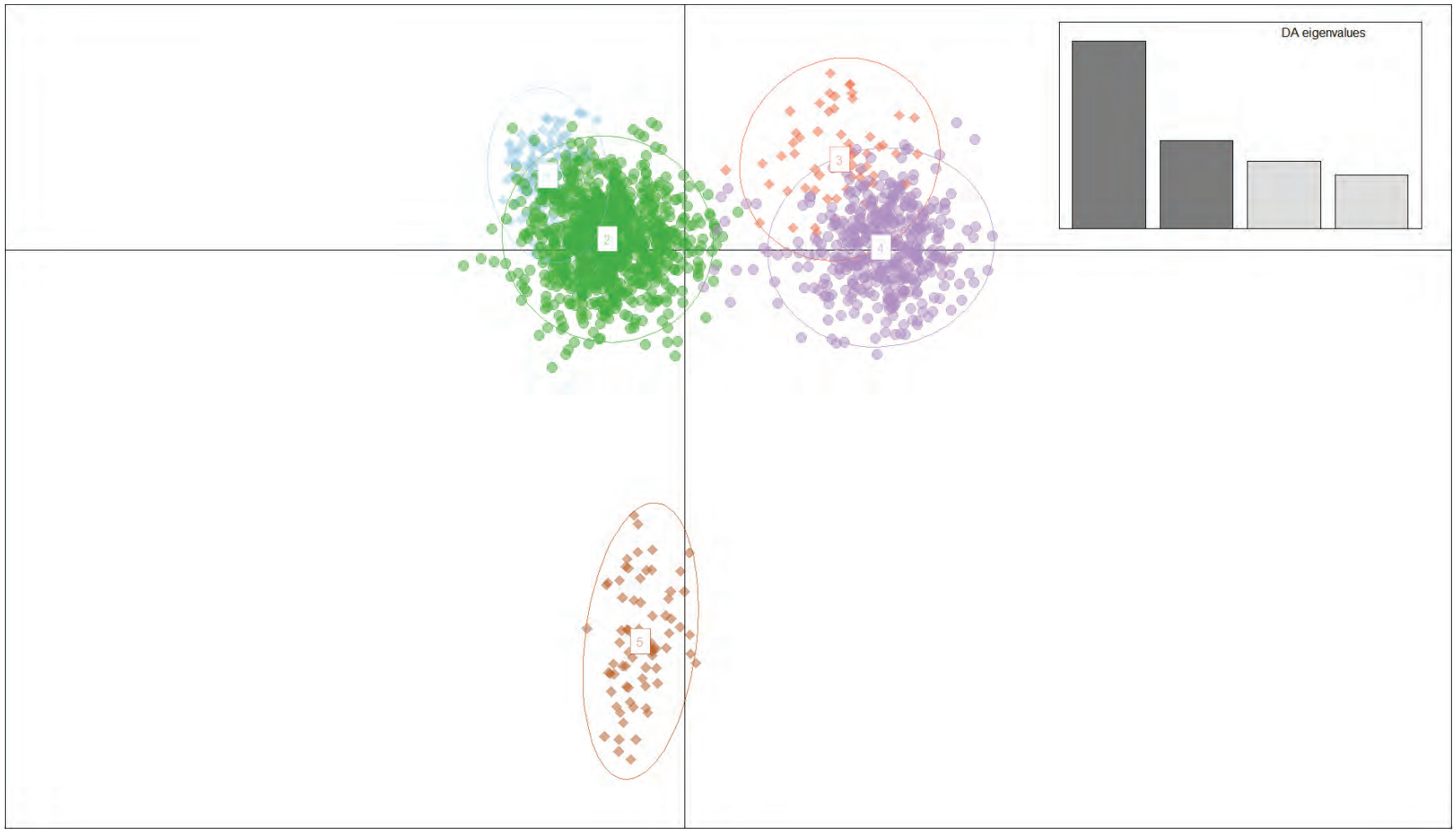


Figure 5.3-13. **Visual representation of DAPC analysis on 1st and 2nd principal components axes of above- and below-Project *O. mykiss* for *k*=5 genetic clusters. First principal component (x axis) pertains to above and below Gorge Dam variance, while 2nd principal component (y axis) was driven by differences at Little Beaver Creek. Cluster 1 was predominantly Three Fools Creek with some Lightning Creek individuals included. Cluster 2 consisted of most of the Project *O. mykiss* populations. Cluster 3 was primarily Pyramid Creek. Individuals from below Project populations resided in cluster 4, with the exception of one individual, and cluster 5 was Little Beaver Creek.**

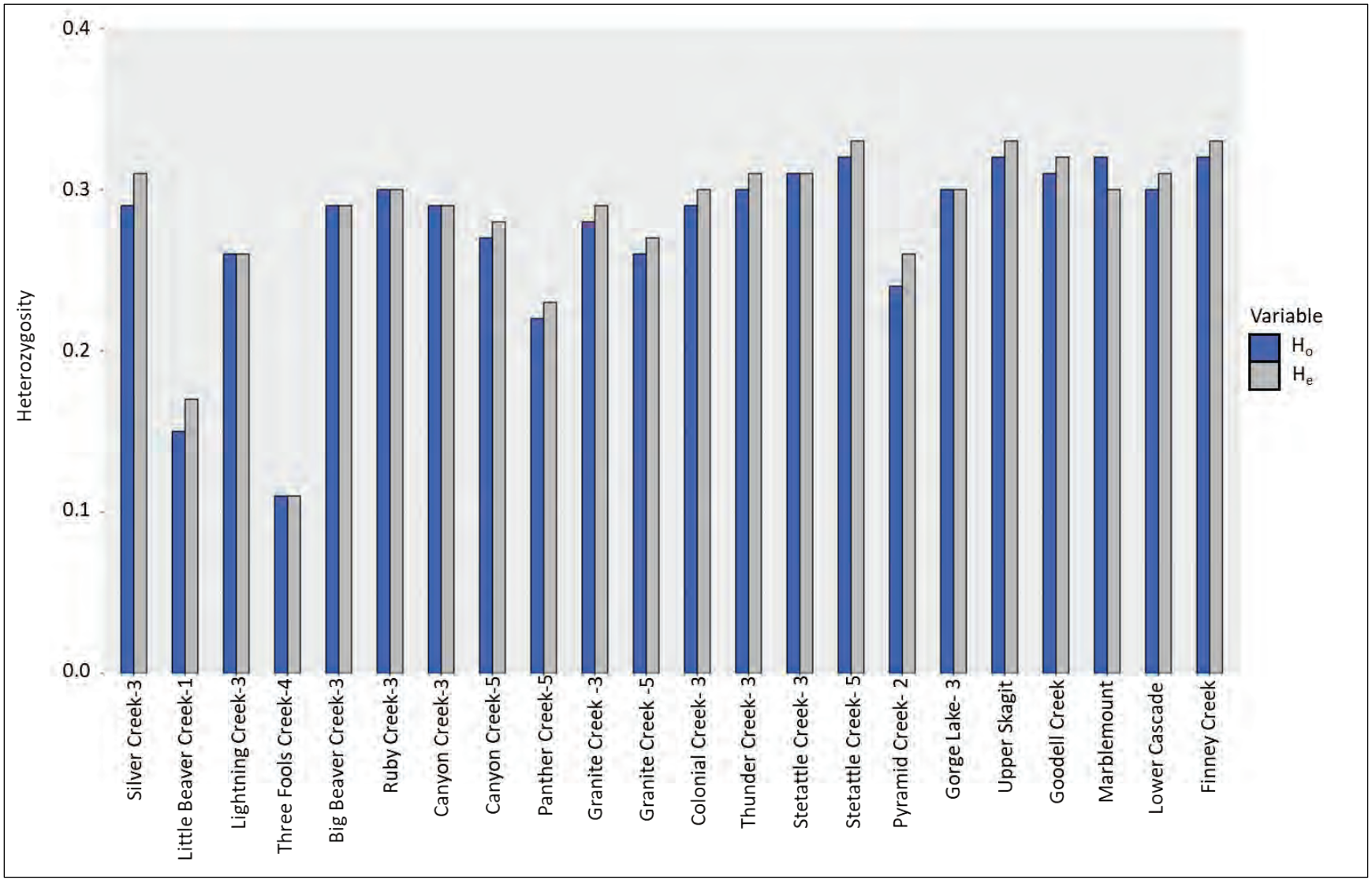


Figure 5.3-14. Genetic diversity (mean observed and mean expected heterozygosity) for above- and below-Project *O. mykiss* populations containing greater than 15 samples.

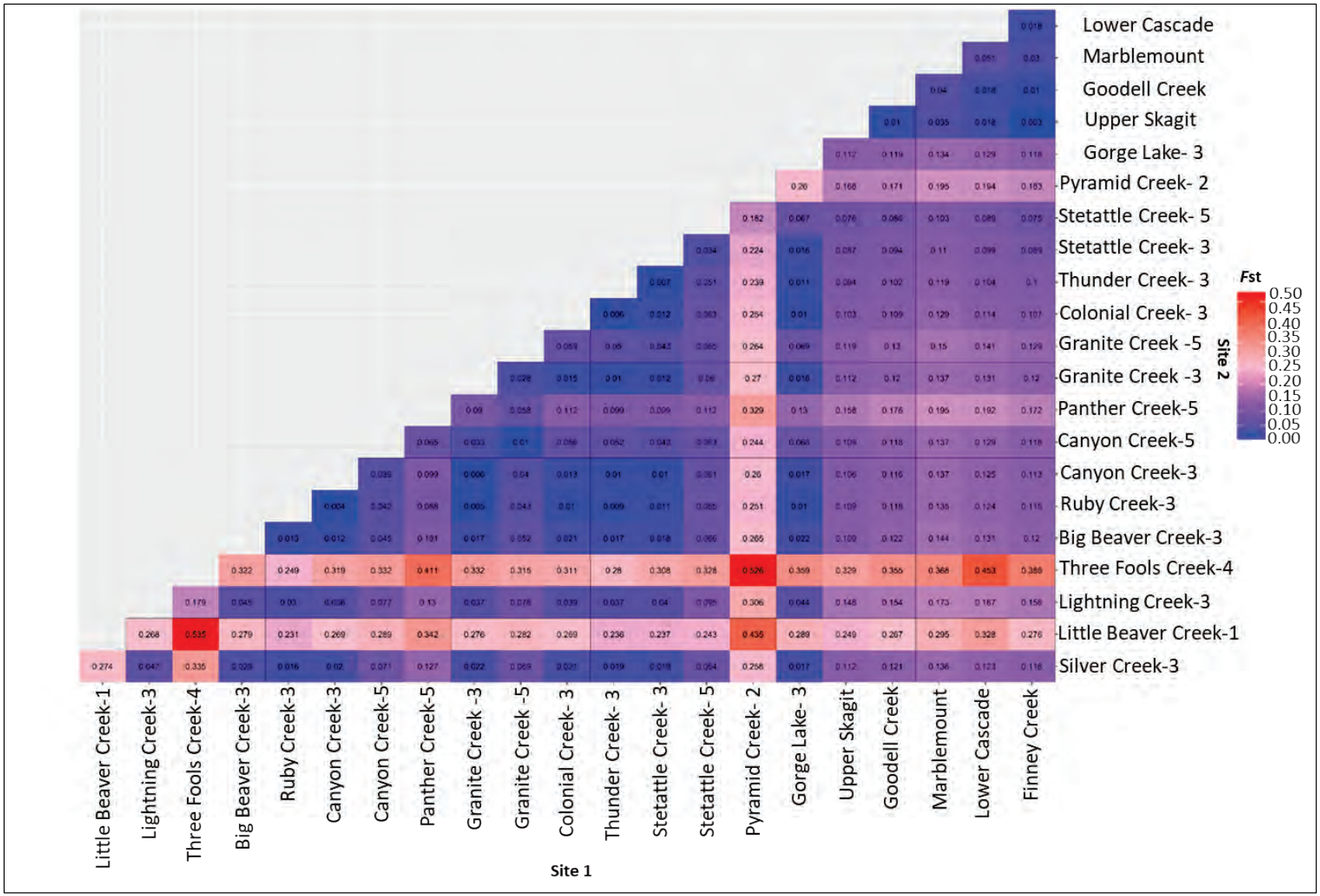


Figure 5.3-15. Pairwise estimates of F_{ST} for above- and below-Gorge Dam O. mykiss populations containing greater than 15 samples. F_{ST} values shown are the actual F_{ST} quantities, and not the significance of each pairwise test.

5.3.1.8 Year 2 Regional Population Analysis

A regional analysis was conducted to provide an assessment of the genetic similarity of Skagit River *O. mykiss* relative to Washington State populations from outside the Skagit River Basin. This objective was accomplished by analyzing 11,653 *O. mykiss* samples represented by 273 populations collections. This total included 30 Project and five WDFW collections from analyses described above. The collections added for this regional analysis were derived from the publicly available Columbia Basin reference genetic baseline (i.e., Columbia River Basin Mykiss GSI baseline v3.3; Hess et. al 2018). There are two caveats to this analysis. First is that while all the 180 genetic loci present in the stock identification reference genetic baseline are included in the 354 SNPs GTSeq panel used by CFS for analysis, approximately 30 percent of these loci are not present in the WDFW population data. Therefore, CFS omitted data (loci) from the total dataset for Project *O. mykiss* to form a complimentary set of data to use with WDFW data for this regional analysis. Second, the Columbia River Basin Mykiss GSI baseline v3.3 does not include data from Salish Sea populations. While these caveats could affect precision of differentiating closely related population aggregates, these data are expected to adequately resolve the primary data pattern between the coastal subspecies of *O. mykiss* (*O. m. irideus*), that is widely distributed along the Western U.S., from populations of the genetically differentiated inland subspecies of *O. mykiss* (redband; *O. m. gairdneri*).

Implementation of the DAPC a-score procedure on the regional dataset containing 11,653 individuals estimated retention of 7 PCs optimized the proportion of successful reassignment corrected for the number of retained PCs (data not shown). DAPC was then rerun on the regional *O. mykiss* dataset using 7 retained PCs, which considered 16.5 percent of the observed genetic variance in the dataset. When considering BIC based selection of various possible number of clusters (*k*), the primary infliction point was unclear, but k=5 (i.e., five genetic clusters; Figure 5.3-16) visualized the primary pattern underlying genetic principal components. Adopting higher k within the DAPC did not change the primary regional relationships, but subdivided populations within the coastal, interior, and study area populations (data not shown). Itemization of where each individual *O. mykiss* resides with respect to membership probability values are not shown here. Broadly speaking about populations present in DAPC analysis, 29 study area populations represented cluster 3 and one study area population (Pyramid Creek) resided in cluster 1. As seen in Figure 5.3-16, the primary axis (x-axis) was driven by variance among coastal *O. mykiss* (*O. m. irideus*) and interior redband (*O. m. gairdneri*). Project populations were placed intermediately along this axis. The secondary axis was driven by variance among Project *O. mykiss* and the 244 populations present in the reference baseline, although one of these 244 populations was Pyramid Creek. Cluster 1 was composed on Pyramid Creek (study area population), five below Project Skagit River populations, one Oregon-Washington Coastal populations, 11 lower Columbia populations, 15 Willamette River populations and 11 middle Columbia River populations (i.e., coastal *O. mykiss*). Cluster 2 comprised 38 lower Snake populations. Cluster 4 was composed of one middle Columbia River and 30 lower Snake River populations. Cluster 5 was composed of 47 middle Columbia River populations, eight upper Columbia populations, 13 Yakima River populations, and 63 lower Snake River populations (i.e., interior *O. mykiss*). Cluster membership for individual Skagit Basin *O. mykiss* is shown in Table 5.3-6. Project *O. mykiss* genetic characteristics appear unique compared to other populations within Washington State.

Table 5.3-6. Counts of regional DAPC cluster membership for Skagit Basin *O. mykiss* samples used for regional genetic analysis. Population labels were retained from Table 5.3-2. Note that no Skagit Basin individuals analyzed assigned to Clusters 2 and 5 (inland redband).

Population	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Silver Creek-1	0	0	1	0	0
Silver Creek-3	0	0	43	1	0
Hozomeen Creek-3	0	0	2	1	0
Little Beaver Creek-5	0	0	73	0	0
Lightning Creek-1	0	0	2	0	0
Lightning Creek-3	0	0	137	1	0
Three Fools Creek-2	0	0	106	0	0
Big Beaver Creek-1	0	0	3	0	0
Big Beaver Creek-3	0	0	49	0	0
McMillan Creek-3	0	0	6	0	0
Pierce Creek-1	0	0	1	0	0
Pierce Creek-3	0	0	2	0	0
Roland Creek-1	0	0	2	0	0
Ross Lake-3	0	0	9	1	0
Ruby Creek-1	0	0	14	0	0
Ruby Creek-3	0	0	170	1	0
Canyon Creek-1	0	0	35	1	0
Canyon Creek-3	0	0	58	0	0
NF Canyon Creek-1	0	0	6	0	0
Panther Creek-1	0	0	17	0	0
Panther Creek-3	0	0	18	0	0
Granite Creek-3	1	0	108	0	0
Colonial Creek-1	0	0	4	0	0
Colonial Creek-3	0	0	49	0	0
Thunder Creek-1	0	0	12	0	0
Thunder Creek-3	0	0	86	0	0
Stetattle Creek-1	3	0	64	2	0
Stetattle Creek-3	0	0	56	0	0
Pyramid Creek-4	59	0	1	0	0
Gorge Lake-3 [1]	0	0	30	0	0
Upper Skagit [1]	147	0	1	0	0
Finney Creek [1]	53	0	0	0	0
Goodell Creek [1]	99	0	0	0	0
Lower Cascade [1]	20	0	0	1	0
Marblemount [1]	106	0	0	0	0

1 WDFW genotype data. These data are from below the Project area (below Gorge Powerhouse).

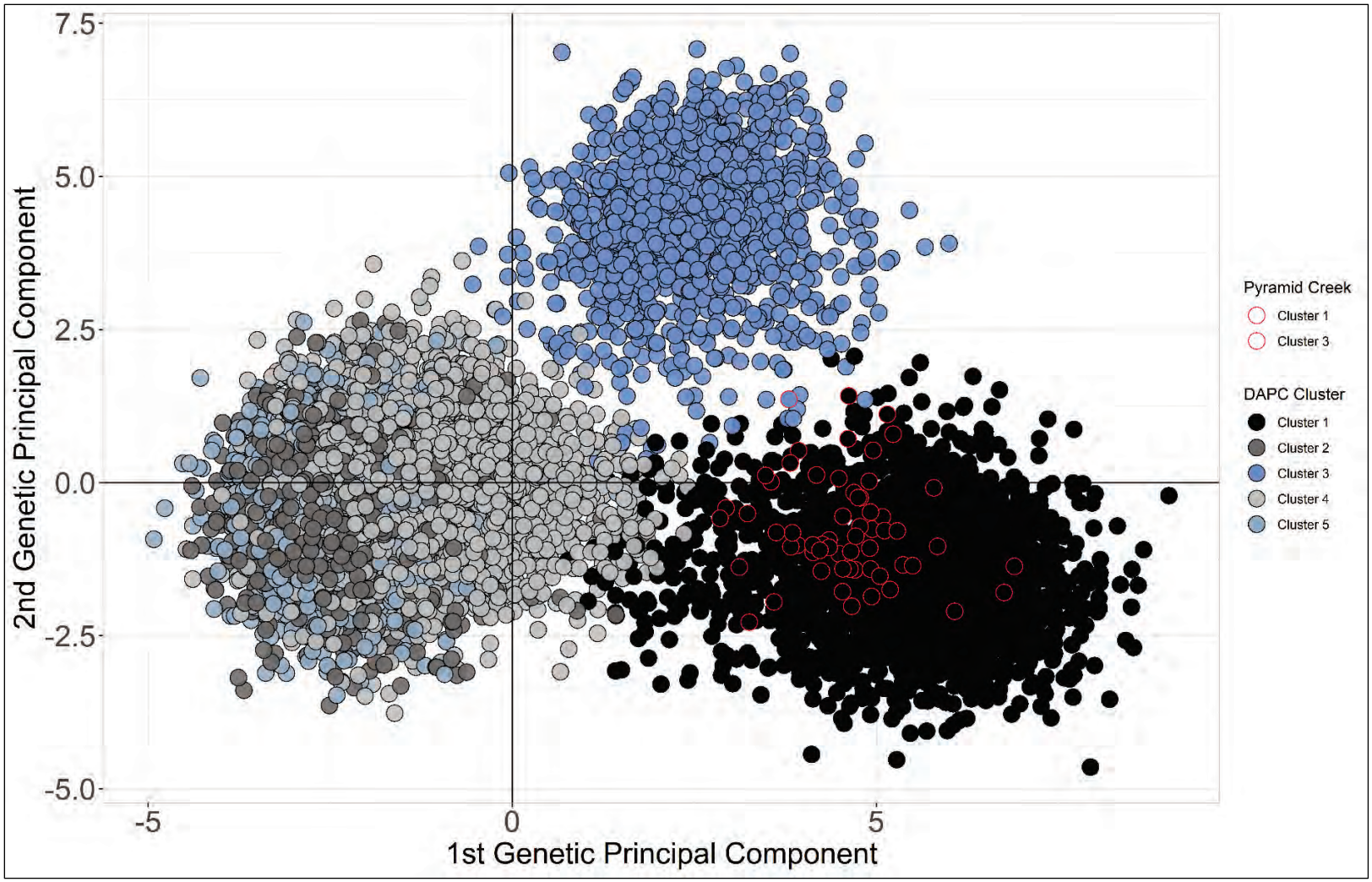


Figure 5.3-16. **Visual representation of DAPC analysis on 1st and 2nd principal components axes of regional *O. mykiss* dataset for *k*=5 genetic clusters. First principal component (x axis) pertains to differences between *O. mykiss* subspecies (coastal versus interior redband), while 2nd principal component (y axis) was driven by differences of Project Area *O. mykiss*. Cluster 3 represents 29 (of 30) Project *O. mykiss* populations. The Pyramid Creek population resides in Cluster 1 with other coastal *O. mykiss* populations (*O. m. irideus*). Clusters 2, 4, and 5 represent inland redband (*O. m. gairdneri*) populations in regional dataset.**

5.3.2 Year 2 Native Char (*Salvelinus* spp.)

5.3.2.1 Year 2 Hybridization

Table 5.3-7 shows the distribution of *Salvelinus* hybrids sampled across the study area in 2022, including samples sent by City Light to CFS. In the total collection size of 374 *Salvelinus* the study team genetically identified 66 Bull Trout (18 percent), 229 Dolly Varden (61 percent), 47 Brook Trout (13 percent), 24 Dolly Varden x Bull Trout hybrids (6 percent), and eight Dolly Varden x Brook Trout hybrids (2 percent). No Bull Trout x Brook Trout hybrids were identified. Hybrids were widely distributed. Specifically, within Ross Lake tributaries, Dolly Varden x Bull Trout hybrids were distributed across 12 sites including Big Beaver (1), Canyon (1), Granite (1), Hozomeen (1), Lightning (2), Roland (2), Ruby (5), Silver (5) (Figure 5.3-17). Four hybrids were detected in Thunder Creek (4) (Diablo Lake tributary), and one hybrid detected at-large from Diablo Lake (J. Fisher City Light). One hybrid was detected from Stetattle (1) (Gorge Lake tributary). For Dolly Varden x Brook Trout, hybrids were detected across five sites including Ross lake tributaries, Pierce (1) and Silver (2) creeks, and Diablo lake tributaries Colonial (1), and Thunder (3) creeks, plus one in a collection from the Gorge reservoir (Fisher 2022). No Bull Trout x Brook Trout hybrids were detected (Figure 5.3-18). Figure 5.3-18 shows a scatterplot of the first two PCs for all *Salvelinus* estimated from all 263 GTseq SNP markers and shows clear distinction among Bull Trout, Dolly Varden, and Brook Trout.

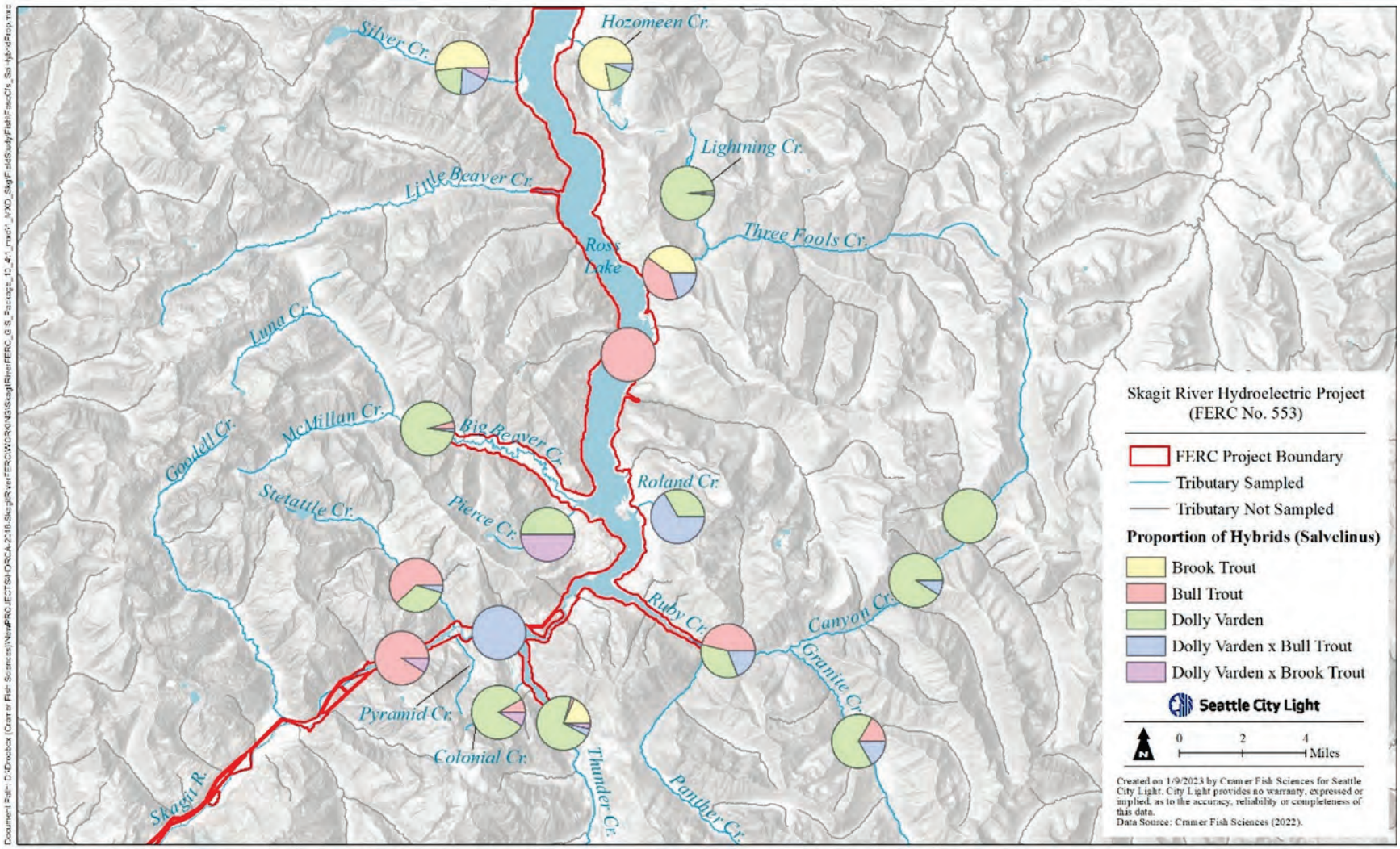


Figure 5.3-17. Map of Year 2 *Salvelinus* collections showing the distribution of the proportion of individuals that were Bull Trout, Dolly Varden, Brook Trout, or hybrids across the Project area based on 22 taxon-diagnostic SNPs.

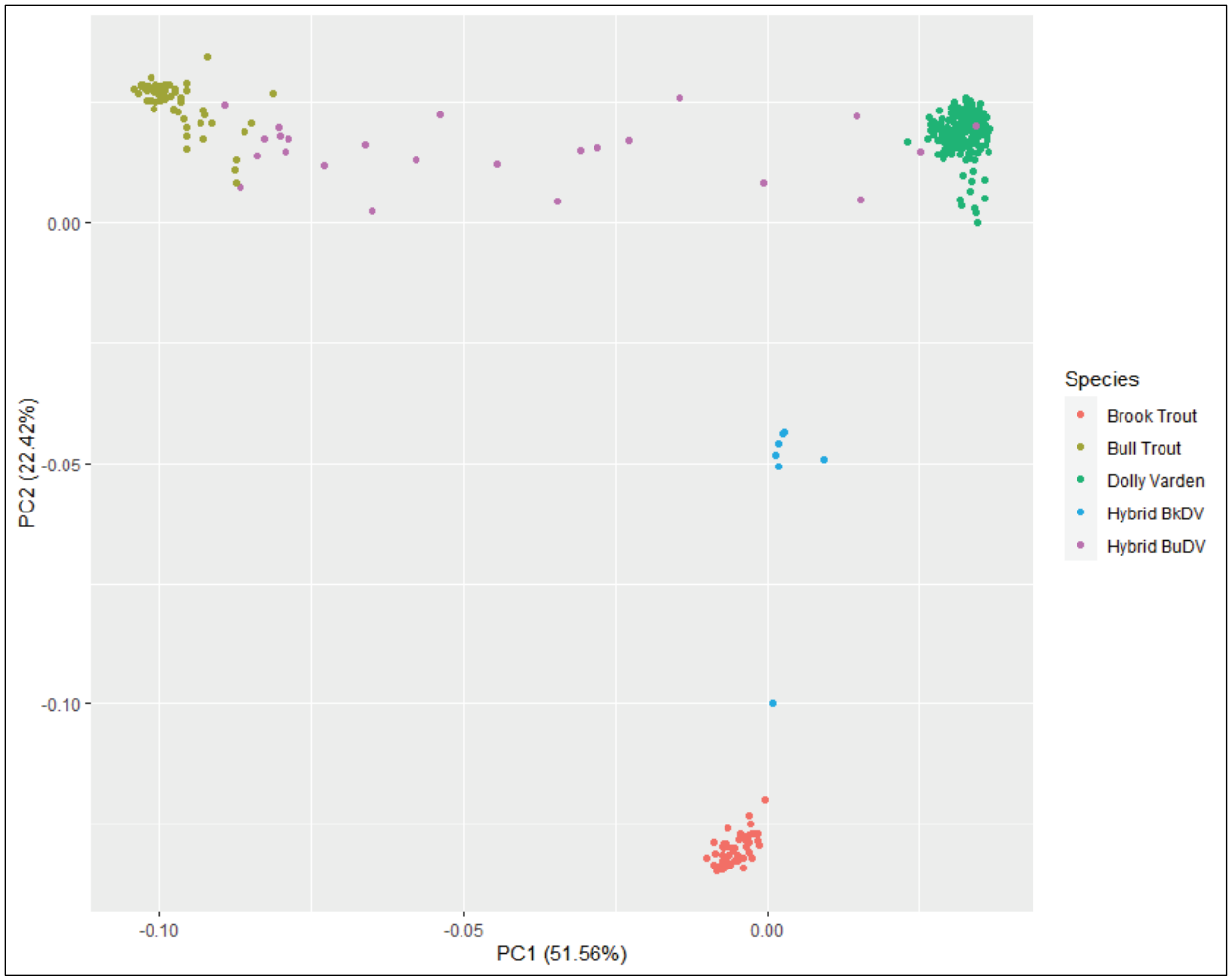


Figure 5.3-18. Scatterplot of first 2 PCs based on genotypes at 263 GTseq SNPs within *Salvelinus* collected in the study area during Year 2 (summer/fall 2022).

5.3.2.2 Genetic variation within collections of Bull Trout and Dolly Varden

A total of 66 Bull Trout were collected from Big Beaver (N=2), Colonial (N=1), Granite (N=1), Ruby (N=12), Stetattle (N=13), Thunder (N=2), the mouth of Lightning Creek (N=2), Gorge Reservoir (N=10), Ross Reservoir (N=22), and mainstem Skagit downstream of the Project (N=1). The use of one Bull Trout from downstream of the Project as a comparison to collections within the study area was due to the fact that genotypes from downstream were not provided by WDFW until after the reporting deadline had passed. To ensure that the study was conducted with the most complete and accurate data available at the time, the researchers chose to use the available Bull Trout genotype from downstream as a comparison to the collections within the study area until the additional genotypes are analyzed. While this was not the ideal scenario, it was necessary to make the best use of the available data. It is important to note that the use of a single genotype as a comparison has limitations and may not be representative of the genetic diversity present in the downstream population.

A key benefit of the GT-seq panel analyzed in year 2 is that it is a standardized panel and thus readily comparable to outside collections that are also genotyped at this panel (e.g., Bohling et al. 2021). Pooling samples from individual tributaries into reservoir-based groups resulted in collections of size 3 in Diablo Lake, 23 in Gorge Lake, and 39 in Ross Lake (Table 5.3-7). From the initial suite of 235 neutral markers, 200 were removed from the analysis due to lack of polymorphism (i.e., Minor Allele Frequency less than 0.01). We removed monomorphic markers, or SNPS with no variability and where all individuals have the same genotype, because they do not contribute variation necessary for statistical analysis. Monomorphic markers are often removed from analyses because they do not provide any information about genetic diversity or differentiation. Nevertheless, when the comparison to samples from downstream of the Project area occurs, the study team will re-evaluate these 200 markers, because relative to other populations, these markers may in fact be quite informative.

When tested within the pooled collections (Ross, Diablo, Gorge), five markers showed significant deviations from Hardy-Weinberg Proportions (HWP) at the α=0.05 level with four tests being significant following sequential Bonferroni correction. Two of the markers (*ScoRAD6812* and *ScoRAD4566*) showed substantial heterozygote excess in Ross Lake (F_{IS}=-0.55 and -0.52, respectively) and Gorge Lake (F_{IS}=-0.89 and -0.60, respectively) and were therefore removed from the analysis. The final Bull Trout dataset contained 33 GTseq SNP markers. H_S was 0.29 (SD=0.16) in Ross Lake and 0.33 (SD=0.15) in Gorge Lake. F_{IS} was 0.03 in Ross Lake and -0.03 in Gorge Lake.

Table 5.3-7. Summary statistics[1] of 2022 Bull Trout (N=65) collections.

Collection pool[1]	N	H_o	H_s	HWP	F_{IS}	N_e
Study Area	65	0.29(0.16)	0.23(0.17)	1/33	-0.26	31.40 (17.50, 68.80)
Ross Lake	39	0.28(0.16)	0.29(0.16)	0/33	0.03	30.9 (17.3, 90.6)
Diablo Lake	3	NA	NA	NA	NA	NA
Gorge Lake	23	0.34(0.17)	0.33(0.15)	1/28	-0.03	10.9 (6.1, 22.4)
inferred1	31	0.32(0.18)	0.31(0.15)	0/30	-0.03	98.50 (26.20, infinite)
inferred2	13	0.43(0.20)	0.38(0.15)	0/24	-0.13	6.7 (2.60, 27.0)
inferred3	21	0.28(0.17)	0.27(0.15)	0/32	-0.04	24.50 (12, 173)

H_S = expected heterozygosity, H_O = observed heterozygosity, HWP = number of markers in significant deviation from Hardy-Weinberg Proportions (α=0.05), F_{IS} = multilocus deviation from expected heterozygosity, N_e = Effective population size (unadjusted mixed cohort). 'NA' indicates sample size was too small to estimate the parameter. Numbers in parentheses are standard deviation except for in the Ne column, in which case parentheses contain the parametric 95 percent CI.

1 Each row represents a distinct collection pool because samples were too small within individual tributaries to be treated separately.

229 Dolly Varden sampled during the 2022 field season were combined with 210 Dolly Varden shared by WDFW for a total of 439. However, 25 individuals were removed from the analysis due to missing genotypes at two or more loci. One additional individual was removed due to an identical genotype in another individual (matching individuals 20NW0447 and 20NW0453 from Lightning Creek, 20NW0447 was retained). The final dataset contained 413 Dolly Varden from

13 tributaries (Table 5.3-8). A total of 102 alleles across the eight microsatellites were observed. Nine of 48 exact tests showed significant deviations from HWP at the α=0.05 level and two were significant after sequential Bonferroni correction. All eight loci were retained because none of the markers consistently deviated from HWP across collections. Mean H_S was 0.65 (SD=0.05) and was lowest in Lightning Creek 0.58 (0.28) and highest in Colonial 0.72 (SD=0.21). Mean F_{IS} across collections was 0.00 (SD=0.04) and ranged from -0.05 in Colonial to 0.04 in Lightning and Ruby. Eighteen of 308 (5 percent) pairwise tests for LD were significant at the α=0.05 level, but none were after correcting for multiple tests.

Table 5.3-8. Summary statistics for collections of Dolly Varden (N=413) and for the 12 inferred genetic groupings identified by DAPC (N=405). Only individuals with greater than 0.50 probability of assignment to an inferred cluster were included (N=405). Note, the twelve inferred clusters are depicted graphically in Figure 5.3-25.

Collection	N	H_o	H_s	HWP	F_{IS}	N_e
Big Beaver	43	0.67 (0.27)	0.67 (0.29)	1/8	-0.01	30.5 (23.4, 41.3)
Canyon	47	0.63 (0.27)	0.64 (0.29)	2/8	0.00	25.1 (19.9, 32.1)
Colonial	22	0.74 (0.17)	0.72 (0.21)	2/8	-0.05	20.7 (13.8, 34.6)
Granite	22	0.64 (0.35)	0.64 (0.35)	0/8	-0.01	24.1 (15.5, 43.5)
Hozomeen	3					
Lightning	136	0.55 (0.26)	0.58 (0.28)	2/8	0.04	21.3 (18.1, 24.9)
NF Canyon	4					
Pierce	1					
Roland	1					
Ruby	29	0.59 (0.32)	0.63 (0.35)	2/8	0.04	21.3 (15.3, 31.6)
Silver	6					
Stetattle	7					
Thunder	92	0.68 (0.28)	0.69 (0.26)	0/8	0.02	34.4 (29.2, 40.7)
Inferred 1	39	0.42(0.26)	0.62(0.23)	3/8	0.27	24.2 (16.9, 36.8)
Inferred 2	43	0.62(0.26)	0.66(0.28)	2/8	0.05	26.0 (20.5, 33.6)
Inferred 3	27	0.67(0.30)	0.65(0.29)	0/8	-0.04	26.2 (18.3, 41.2)
Inferred 4	27	0.67(0.29)	0.65(0.27)	0/8	-0.03	17.8 (12.0, 28.2)
Inferred 5	37	0.62(0.26)	0.67(0.28)	3/8	0.06	33.5 (25.6, 46.0)
Inferred 6	38	0.64(0.32)	0.64(0.32)	0/8	-0.01	28.7 (22.0, 38.7)
Inferred 7	33	0.67(0.34)	0.66(0.31)	0/8	0.01	36.2 (25.9, 54.7)
Inferred 8	41	0.61(0.27)	0.64(0.29)	1/8	0.05	31.8 (23.7, 44.8)
Inferred 9	36	0.68(0.26)	0.67(0.28)	1/8	-0.02	28.3 (21.5, 38.6)
Inferred 10	43	0.62(0.32)	0.64(0.32)	1/8	0.03	34.2 (26.0, 47.0)
Inferred 11	15	0.64)0.32)	0.66(0.30)	0/8	0.03	20.1 (11.4, 48.5)
Inferred 12	26	0.71(0.23)	0.69(0.23)	1/8	-0.05	31.0 (20.9, 52.3)

Notes: H_S = expected heterozygosity; H_O = observed heterozygosity; HWP = number of markers in significant deviation from Hardy-Weinberg Proportions (α=0.05); F_{IS} = multilocus deviation from expected heterozygosity; N_e = Effective population size (unadjusted mixed cohort).

'NA' indicates sample size was too small to estimate the parameter. Numbers in parentheses are standard deviation except for the N_e column where parentheses contain the parametric 95 percent CI.

5.3.2.3 Year 2 Genetic divergence among collections of Bull Trout and Dolly Varden

To explore divergence among Bull Trout, samples were initially grouped by their reservoir of origin due to small sample sizes within individual tributaries. AMOVA based on reservoir groupings showed that reservoirs account for 5 percent of genetic variation (F_{ST}=0.05; P<0.01) and that variation within individuals accounts for the remaining 95 percent. Mean pairwise F_{ST} between reservoirs was F_{ST} = 0.03 (95 percent CI: 0.02 to 0.05) (Table 5.3-9). The highest divergence occurred between Diablo and Gorge lakes (F_{ST}=0.05) and the lowest between Gorge and Ross lakes (F_{ST}=0.03). By contrast, sample sizes tended to be large enough within Dolly Varden to analyze them by tributary. Mean pairwise F_{ST} between collections of Dolly Varden was F_{ST}=0.05 (95 percent CI: 0.037, 0.055). The highest divergence occurred between Lightning Creek and Roland Creek (F_{ST}=0.16). The lowest divergence occurred between Silver and Stetattle Creek (F_{ST}=-0.03), a negative result likely due to small sample size (N=6 and 7, respectively). The next-lowest divergence occurred between Ruby Creek and Canyon Creek (F_{ST}=0.003), which was not unexpected, given Canyon and Ruby Creeks are only nominally distinct (Canyon Creek becomes Ruby Creek in the lower watershed). Unlike Bull Trout, hierarchical AMOVA based on reservoir groupings (i.e., nesting tributaries within reservoirs) did not explain a significant amount of genetic divergence in Dolly Varden (F_{CT}=0.006; P=0.32). However, the Mantel tests conducted in this study show a strong relationship between geographic and genetic distance (P<0.01; R^2=0.40), indicating that Dolly Varden is genetically structured by isolation-by-distance. However, the positive relationship between geographic distance and the scatter of residuals from the isolation-by-distance analysis was not significant (P=0.11, R^2=0.10). Upon reanalyzing the data without an outlier, the Mantel test remained non-significant, and the strength of the relationship did not change substantially (P=0.09; R^2=0.09). While the pattern of isolation by distance is consistent with equilibrium, the scatter of residuals is not. Therefore, the presence of equilibrium is somewhat ambiguous. This suggests that while the genetic structure of Dolly Varden is influenced by geographic distance, there may be other factors affecting gene flow and drift such that a significant relationship between geographic distance and genetic structure is not present.

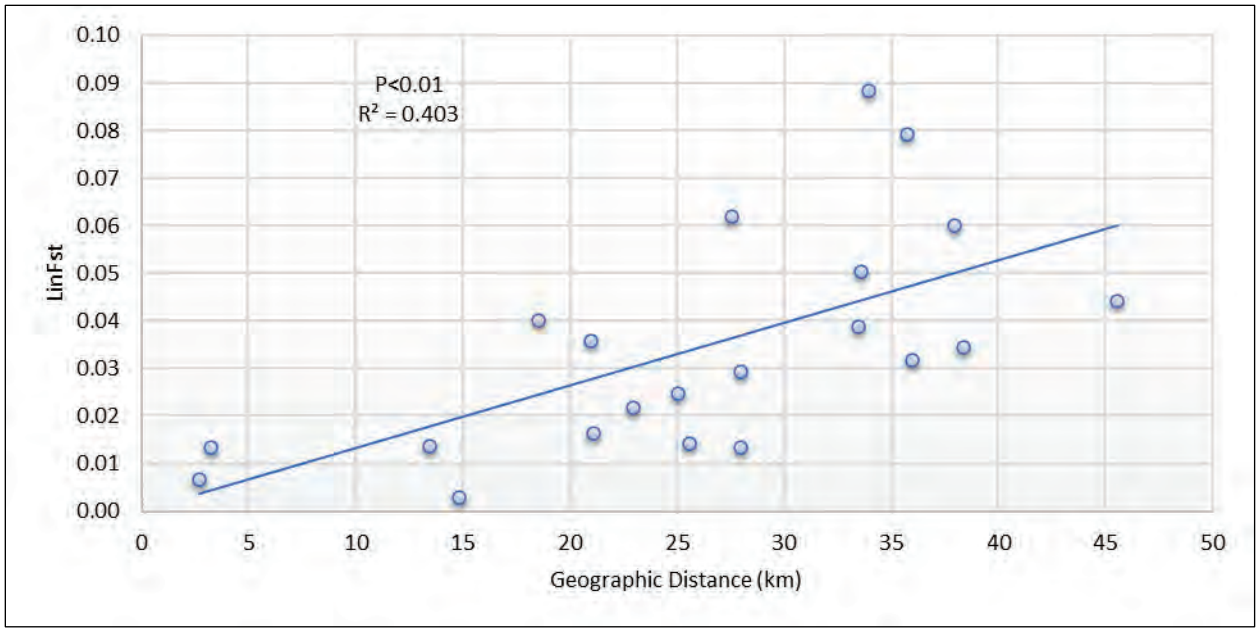


Figure 5.3-19. **Isolation by distance analyses for Dolly Varden in the study area assayed at 8 microsatellite loci. Linear pairwise F_{ST} distances are plotted against pairwise geographical distance. The Mantel test suggests that 40 percent of the variability observed in the F_{ST} is explained by geographic distance.**

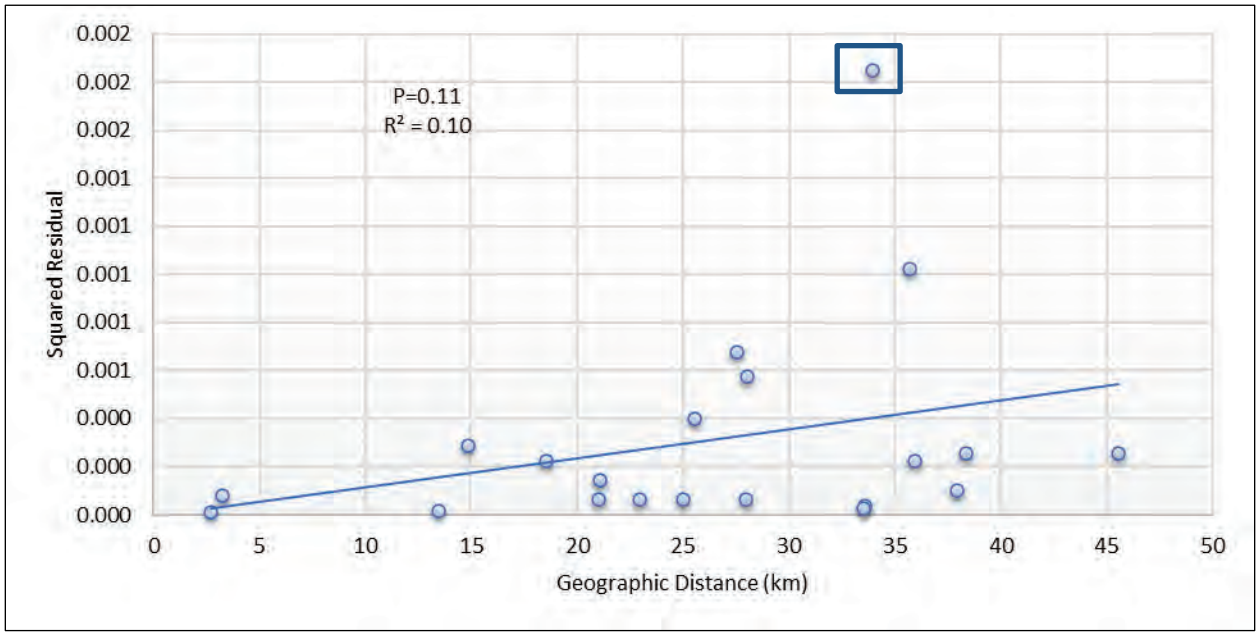


Figure 5.3-20. **Scatterplot of the squared residuals from the isolation-by-distance analysis in Dolly Varden using 8 microsatellite loci. Although a positive relationship was observed (R^2=0.10), the Mantel test was not statistically significant (P=0.11). The analysis was rerun without the apparent outlier (Point in the box, Colonial and Lightning) and the interpretation did not change (i.e., the Mantel test remained nonsignificant).**

Table 5.3-9. Pairwise F_{ST} for Bull Trout collection pools based on reservoirs and inferred genetic clusters from DAPC analysis. Statistically significance estimates (at alpha=0.05 level) are indicated by bold lettering.

	Diablo Lake	Gorge Lake	Ross Lake	inferred1	inferred2	inferred3
Diablo Lake	0					
Gorge Lake	0.04729	0				
Ross Lake	0.02634	**0.02623**	0			
inferred1	0.0052	**0.02554**	**0.0226**	0		
inferred2	**0.13765**	**0.04016**	**0.09551**	**0.14063**	0	
inferred3	**0.09041**	**0.08098**	0.01246	**0.09945**	**0.15103**	0

5.3.2.4 Year 2 Discriminant Analysis of Principal Components

A scatterplot of the first two PCs appeared to show genetic structuring (Figure5.3-21) of Bull Trout in the study area, yet specific patterns were visually obscure (i.e., not obviously associated with contemporary watershed boundaries). DAPC was implemented to identify and describe clusters of genetically related individuals. The most optimal *k*-means-based clustering solution occurred when k=3 (BIC=98.43), but we also explored k=2 (BIC=99.63) and k=4 (BIC=98.51) due to comparable model support. The *k*-means clustering algorithm using all 31 principal components (i.e., the "uninformed prior" population assignments) did not appear to group Bull Trout into clusters that conformed strongly with obvious contemporary geographic features, such as reservoirs or tributaries. One possible exception was that the three Bull Trout collected in Diablo Lake, all clustered together for k=2 and k=3, albeit with fish from both Gorge and Ross lakes.

Discriminant analysis of the prior inferred k-means based clusters using the first 20 principal components explained 94.5 percent of the variance for k=2 and k=3 (1 and 2 discriminant functions, respectively) and 86.1 percent of the variance for k=4 (3 discriminant functions). Scatterplots of the discriminant functions clearly distinguished the inferred genetic clusters visually with little to no overlap among clusters (Figure 5.3-22). The posterior probability of assignment back to the prior inferred clusters was 1.00 for all *k*, suggesting clear-cut genetic groups exist in the study area. Nevertheless, the a-score (an index of overfitting) suggested that 20 principal components was likely an overfit of the data and so the discriminant analysis was rerun for k=3 (i.e., the most supported model) using a more optimal number of 6 PCs. Posterior assignments were not associated with any apparent contemporary physical features, such as reservoirs or tributaries. The scatterplot for k=3 did not consistently group Bull Trout into collections based on current reservoir boundaries, suggesting contemporary reservoirs may not provide a complete picture of the genetic structure of Bull Trout in the study area (Figure 5.3-23). Specifically, the most supported model placed Bull Trout into three genetic clusters that were well mixed between Ross Lake and Gorge Lake.

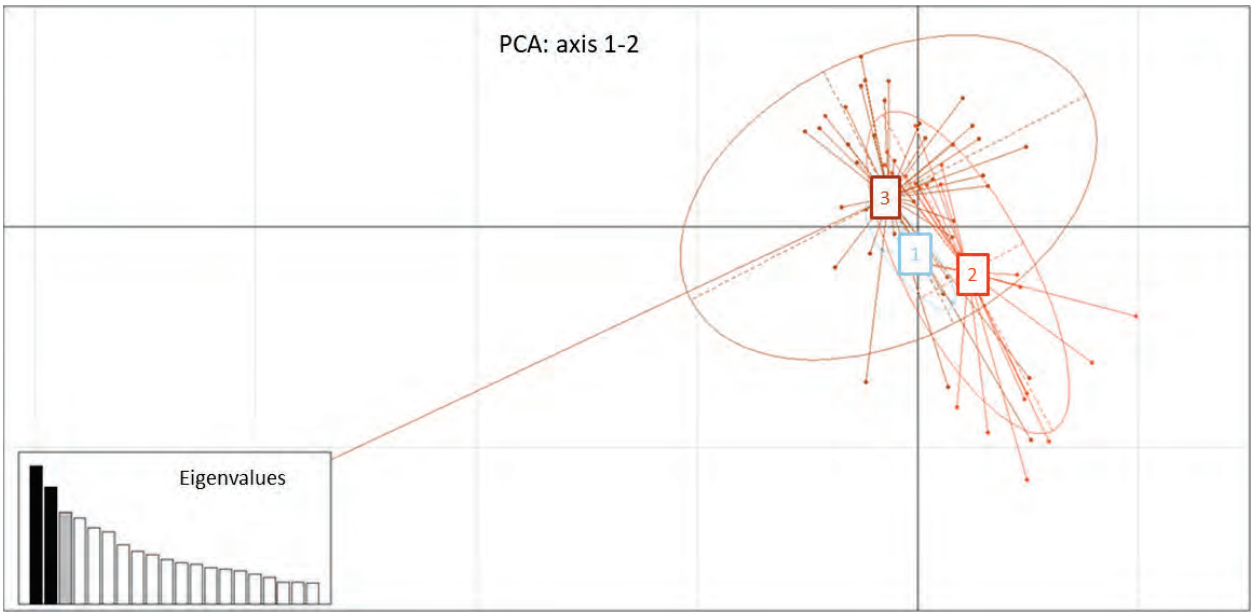


Figure 5.3-21. **Scatterplot of the first 2 PCs based on 33 GT-Seq SNP genotypes in 65 Bull Trout sampled in Year 1 (2022) Diablo (centroid 1) Gorge (centroid 2) Ross (centroid 3) lakes. From the initial suite of 235 neutral markers, 200 were removed from the analysis due to lack of polymorphism (i.e., Minor Allele Frequency less than 0.01). Bull Trout were grouped by Reservoir because sample sizes were too small to analyze by tributaries.**

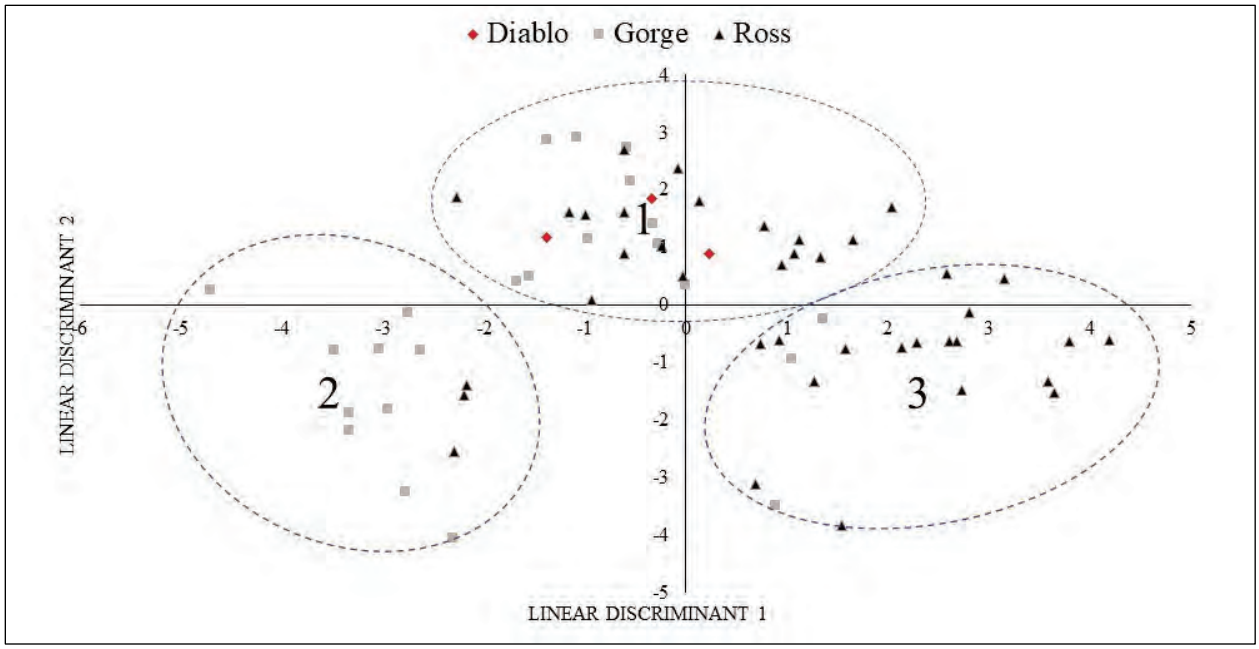


Figure 5.3-22. **Scatter plot of the first two linear discriminants produced by the final DAPC for 65 Bull Trout that were sampled during the 2022 field season (year 2) and genotyped using 33 GT-seq SNP markers. The analysis inferred three genetic clusters, which are identified by the numbers (in black) in each ellipse. The color and shape of each point, as shown in the legend, indicates whether the Bull Trout was sampled in Diablo, Gorge, or Ross Reservoirs. Genetic cluster 1 contained Bull Trout from all three reservoirs. Nevertheless, Bull Trout from Ross tended to have higher loadings for the first linear discriminant, while those from Gorge tended to have lower loadings. All three Bull Trout from Diablo were grouped within cluster 1, although this sample size was very small. The analysis used k=3, 6 Principal Components (PCs), and two discriminant functions. The specific driver of the genetic structure observed is unclear, but 100 percent accuracy of posterior assignments back to each cluster suggests genetic structure among Bull Trout is present.**

For Dolly Varden, the first two PCs showed genetic structuring, but specific patterns or drivers were not visually obvious (Figure 5.3-24). Considering all 101 PCs in the DAPC, the most optimal k-means-based clustering solution for Dolly Varden occurred when k=12 (BIC=339.01). Discriminant analysis of these 12 clusters using the first 40 principal components and 11 discriminant functions explained 94 percent of the variance for k=12. Visually, a scatterplot of the first two discriminant functions showed substantial overlap among the inferred genetic clusters (Figure 5.3-25). Nevertheless, the posterior probability of assignment of individuals back to the 12 inferred clusters was very high (95 percent accurate), suggesting the 12 groups reflect a tangible and substantive underlying genetic structure in Dolly Varden. Nevertheless, the a-score suggested 40 principal components likely provides an overfit of the model and so the discriminant analysis was rerun for k=12 (i.e., the most supported model) using a more optimal number of 15 PCs. The optimized model performed nearly as well, providing 94 percent accuracy of posterior assignments back to inferred clusters. Intriguingly, composition plots of the posterior assignments to the 12 genetic clusters consistently grouped individuals from very distal watersheds together, again

highlighting a contradictory pattern relative to the isolation-by-distance analysis that suggested neighboring populations should consistently contain relatively similar allele frequencies.

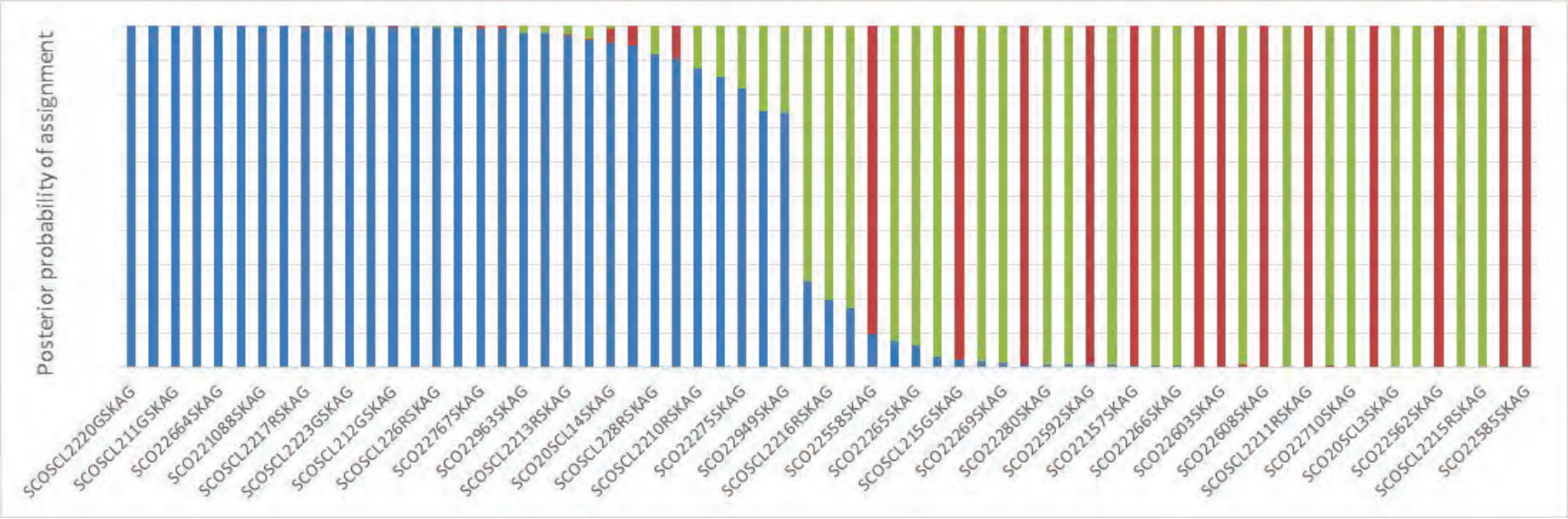


Figure 5.3-23. **Composition plot that displays the posterior probability of assignment of 65 Bull Trout samples from the Project area to k=3 inferred DAPC clusters during year 2. Each vertical line in the plot represents an individual fish, and the color of the line represents its posterior probability of assignment to three inferred genetic clusters. The results indicate that Bull Trout did not cluster entirely by reservoir. The individuals are sorted based on their posterior probability of assignment to each of the three inferred clusters identified using DAPC.**

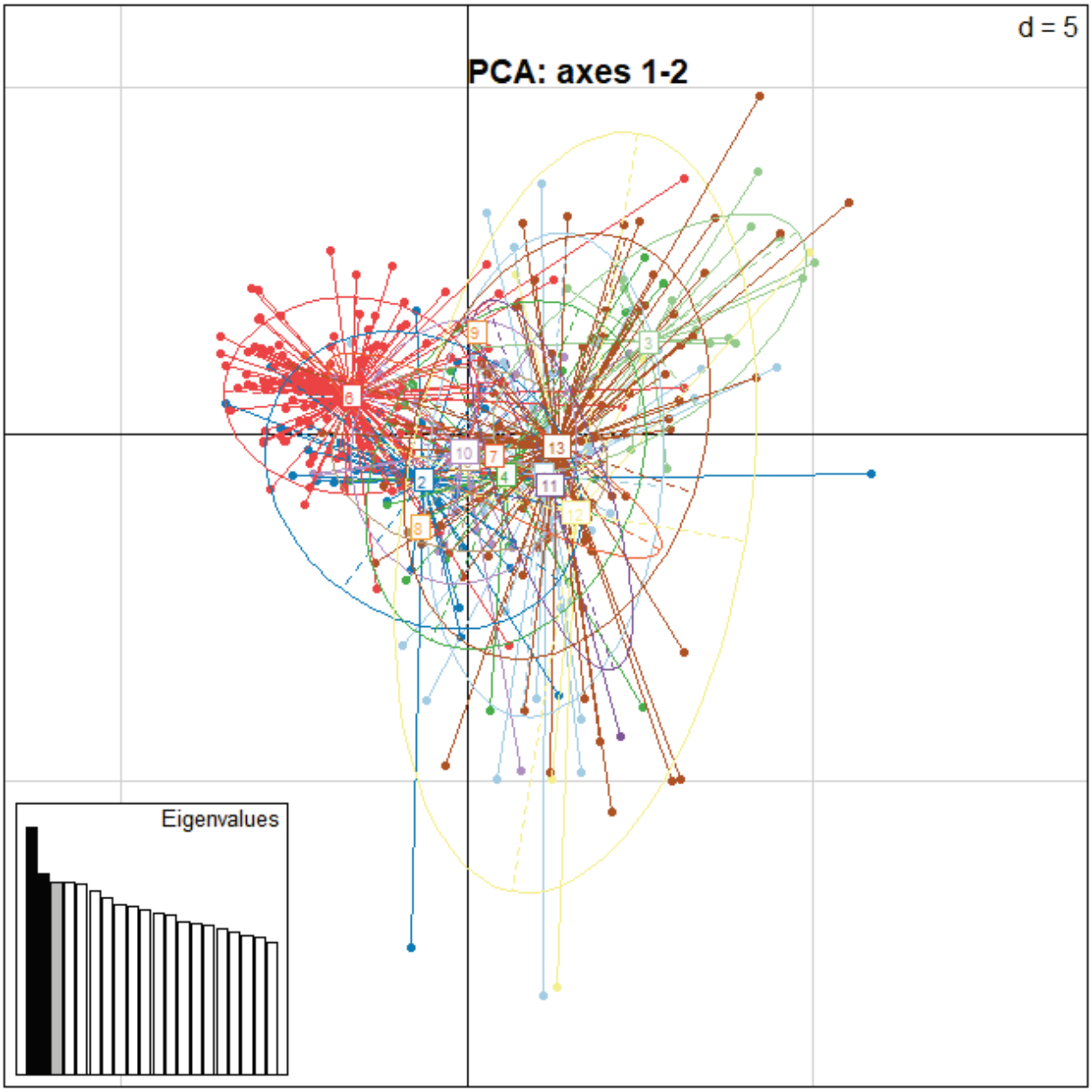
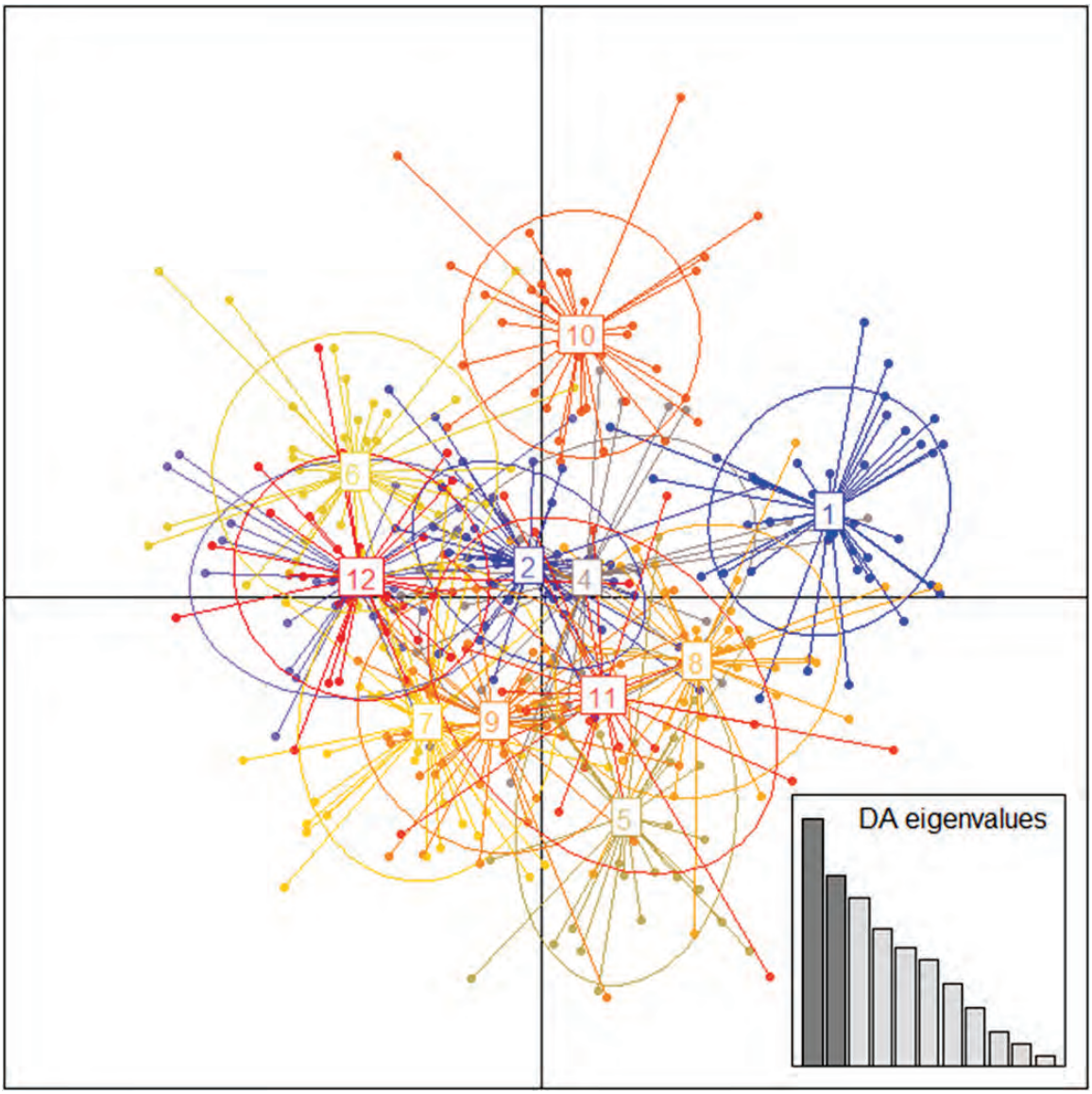
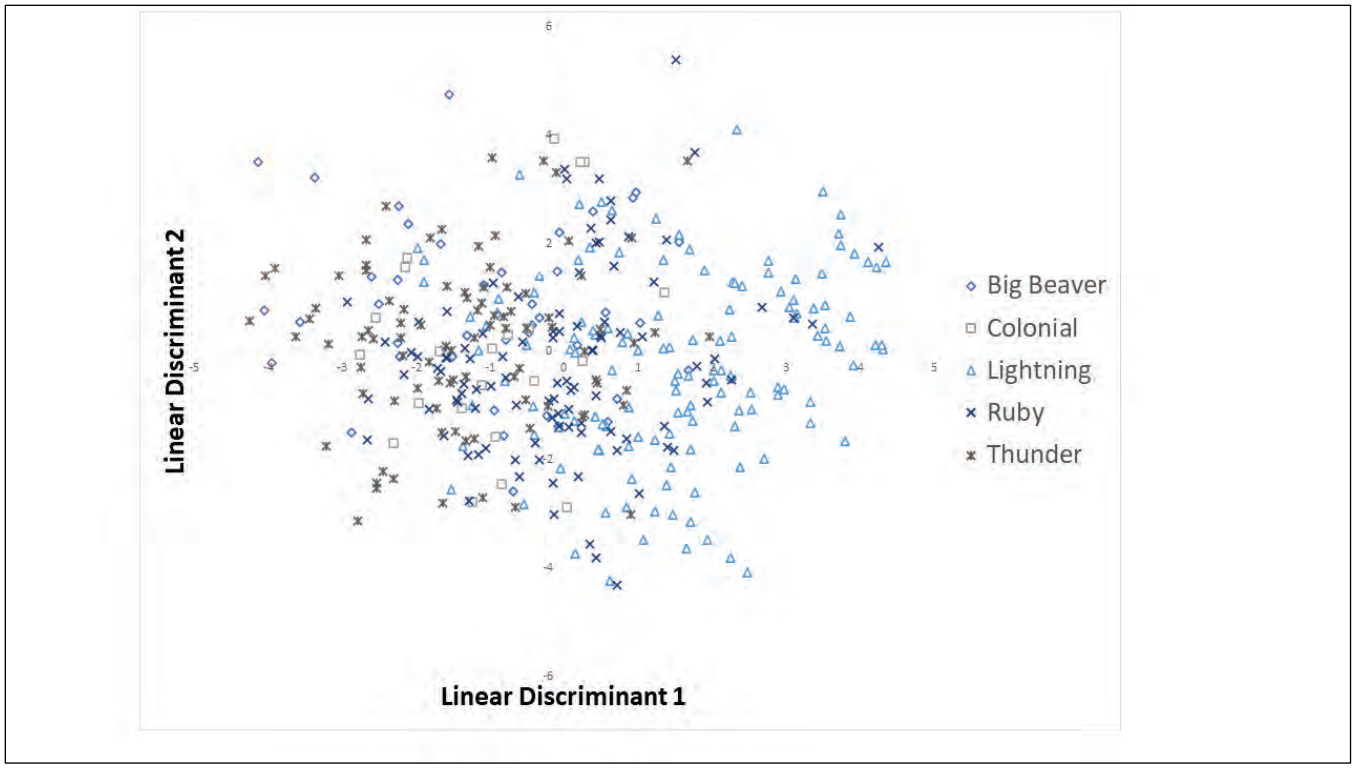


Figure 5.3-24. **Scatterplot of first 2 PCs based on genotypes of 413 Dolly Varden sampled in 13 collections during year 2. Analysis is based on genotypes at eight microsatellites. Each point represents an individual Dolly Varden. Note: (1) Big Beaver, (2) Canyon, (3) Colonial, (4) Granite, (5) Hozomeen, (6) Lightning, (7) NF Canyon, (8) Pierce, (9) Roland, (10) Ruby, (11) Silver, (12) Stetattle, (13) Thunder creeks.**



Top panel Figure 5.3-25. This panel is a scatterplot of the same data presented in the bottom panel; it is just grouped by the 12 inferred clusters identified by DAPC instead of the tributary in which samples were collected. Inferred clusters do not necessarily correspond to specific tributaries because the analysis is "unsupervised."

Figure 5.3-25. **Page 1 of 2. The scatterplot shows the discriminant function scores for seven microsatellites that were genotyped in Dolly Varden. The analysis assumes K=12 and uses 40 principal components. Each individual is represented by a point in the scatterplot. The inset in this panel displays the eigenvalues of the analysis. The plot displays the projection of the first two linear discriminants. The x-axis represents the first linear discriminant, and the y-axis represents the second linear discriminant. Two panels are presented to facilitate interpretation of the results. In this panel, elipses are associated with the 12 inferred clusters. In the next panel, individuals are colored and shaped according to their tributary of origin. Both panels show the exact same information the points are just colored/shaped differently depending on whether they are displaying the inferred clusters (top panel) or tributary of origin (bottom panel)**



Bottom panel Figure 5.3-25. This panel is a scatterplot of the same data shown in the top panel, it is just not grouped by the 12 inferred clusters. Instead, it is colored/identified based on the tributary the samples were collected in.

Figure 5.3-25. **Page 2 of 2. In the previous page, the scatterplot shows the discriminant function scores for seven microsatellites that were genotyped in Dolly Varden. The analysis assumes K=12 and uses 40 principal components. Each individual is represented by a point in the scatterplot. The inset in the previous panel displays the eigenvalues of the analysis. The plot displays the projection of the first two linear discriminants. The x-axis represents the first linear discriminant, and the y-axis represents the second linear discriminant. Two panels are presented to facilitate interpretation of the results. In the previous panel, elipses are associated with the 12 inferred clusters. In this panel, individuals are colored and shaped according to their tributary of origin. Both panels show the exact same information the points are just colored/shaped differently depending on whether they are displaying the inferred clusters (top panel) or tributary of origin (bottom panel)**

5.3.2.5 Year 2 Scale Age Determination for Estimating N_e and N_b

Scales were collected to estimate ages of fish in support of estimating effective population size. During the 2022 field season, scales were collected from 255 *Salvelinus* individuals of which 110 were aged. The 110 *Salvelinus* samples contained four age classes (0+, 1+, 2+, 3+, and 5+) with a range of 1-60 individuals per age class (Table 5.3-10). Age class 0+ and 1+ displayed the most overlap of fork lengths between age classes while other age classes displayed little overlap (Figure 5.3-26).

Table 5.3-10. Size and age summaries for scale age determined *Salvelinus* individuals.

Age Class	n	Min	Max	Mean	SD
0+	33	60	118	78.97	12.95
1+	60	76	188	123.41	21.4
2+	9	160	256	219.18	24.2
3+	7	323	426	356.25	35.31
5+	1	420	420	420	0

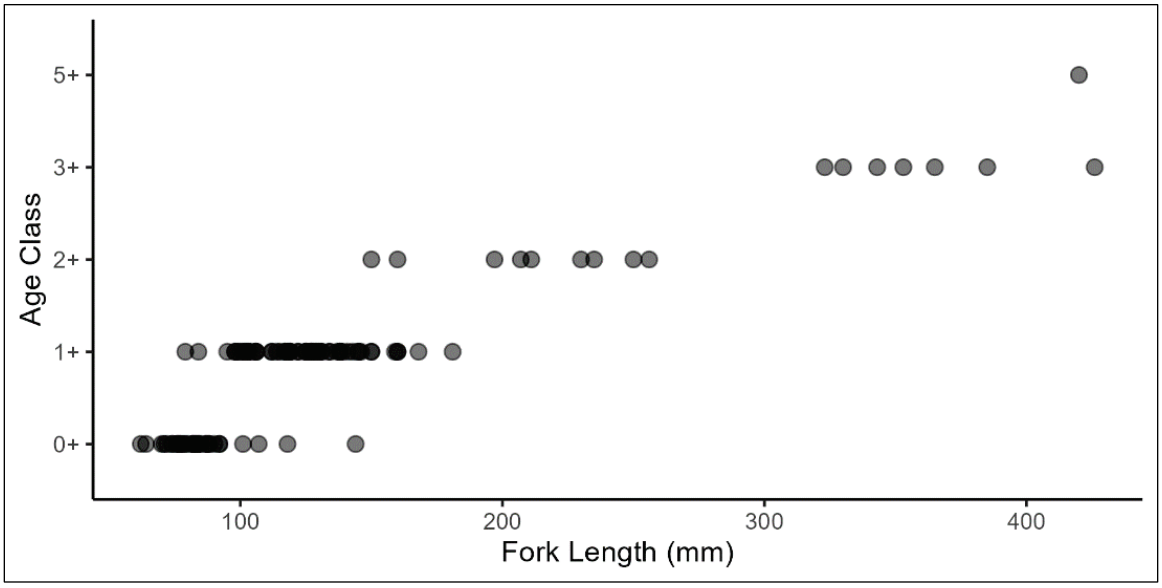


Figure 5.3-26. Summary of age, based on scale analysis, and fork length (mm) of 110 *Salvelinus* captured in 2022.

5.3.2.6 Year 2 Length-at-Age Key Assignment

A total of 201 un-aged *Salvelinus* samples were assigned ages based on the length-at-age key for a total of 311 individuals and had a range of 2-175 individuals per age class (Table 5.3-11). Age 0+ fork lengths ranged from 60–118 mm, age 1+ fork lengths ranged from 76–188 mm, age 2+ fork lengths ranged from 160–256 mm, age 3+ fork lengths ranged from 323-426 mm, and age 5+ had two individuals with fork lengths 420 mm. There was considerable overlap of fork lengths between age classes 0+, and 1+ (Figure 5.3-27).

Table 5.3-11. Size and age summaries for aged and age-assigned *Salvelinus* individuals.

Age Class	*n*	Min	Max	Mean	SD
0+	98	60	118	78.97	12.95
1+	175	76	188	123.41	21.4
2+	28	160	256	219.18	24.2
3+	8	323	426	356.25	35.31
5+	2	420	420	420	0

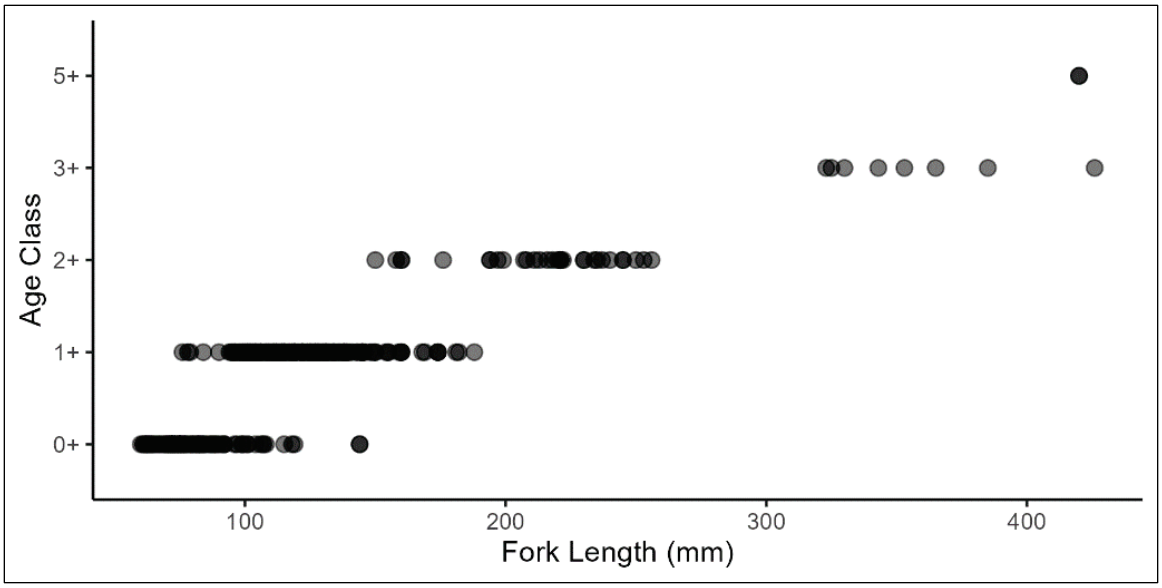


Figure 5.3-27. Summary of age, based on scale analysis and age-assignment, and fork length (mm) of 311 *Salvelinus*.

5.3.2.7 Year 2 Effective population size

Tables 5.3-7, 5.3-8, 5.3-12, and 5.3-13 contain summaries of effective size estimates for Bull Trout and Dolly Varden. N_e for Bull Trout was 31.40 (95 percent CI 17.50, 68.80) when all 65 individuals were analyzed as a collection from a single population (i.e., uncorrected mixed-cohort N_e). When samples were divided into groups based on their sampling location (i.e., Ross, Diablo, or Gorge Lake), N_e was 30.9 (95 percent CI 17.3, 90.6) in Ross Lake, 10.9 (95 percent CI 6.1, 22.4) in Gorge Lake, and was inestimable (-1.2) in Diablo Lake due to small sample size (N=3). When samples were divided into the three inferred genetic clusters identified by the DAPC, the effective sizes were 98.50 (95 percent CI 26.20, infinite) for inferred cluster 1 (k1), 6.7 (95 percent CI 2.60, 27.0) for inferred cluster 2, and 24.50 (95 percent CI 12, 173) for inferred cluster 3. The effective number of breeders (N_b) was also attempted to be estimated by grouping individuals into cohorts. Sample sizes within individual cohorts were too small, however, so fish were grouped into two mixed cohort groups: one group consisted of age-0 to 1+ (N=29) and the second group consisted of 2+ (N=36). Raw N_b for age 0 to 1+ was 6.4 (95 percent CI 3.2, 10.4) and the N_b for age 2+ was 47.2 (95 percent CI 23.4, 194). These raw N_b estimates were used to calculate adjusted N_e using Waples et al. (2014) "two trait" correction formula, which produced corrected estimates 12.08 and 88.35 for group 1 and group 2, respectively (Table 5.3-12).

For Dolly Varden, uncorrected, mixed cohort N_e was 24.53 (harmonic mean [95 percent CI 18.19, 34.36) and ranged from 20.7 (95 percent CI 13.8, 34.6) in Colonial Creek to 30.5 (95 percent CI 23.4, 41.3) in Big Beaver Creek. When collections were divided into the 12 inferred clusters based on DAPC, harmonic mean N_e was 27.00 (95 percent CI 19.52, 41.10) and ranged from 17.80 (95 percent CI 12.0, 28.2) in cluster 4 to 36.20 (95 percent CI 25.9, 54.7) in cluster 7. Adjusted N_b was 20.61 (harmonic mean [95 percent CI 14.47, 31.24) and ranged from 13.02 (95 percent CI 8.94, 19.90) in Lightning Creek in 2021 to 30.42 in Thunder Creek in 2021. Adjusted N_e was 42.93

(harmonic mean [95 percent CI 26.87, 76.20]) and ranged from 27.24 (95 percent CI 21.02, 45.97) in Lightning Creek in 2021 to 63.13 (95 percent CI 55.24, 91.33) in Thunder Creek in 2021.

Table 5.3-12. Effective population size estimates for Bull Trout corrected for overlapping generations using Waples et al. (2014) adjustment based on adult life span (8.5 years) and age of first reproduction (3.0) (Hemmingsen et al. 2001). Bias corrections are insensitive to Adult Lifespan and Age of Maturity within a few years. Only collections with greater than 20 samples were analyzed.

Population	Cohort	Age Class	N_b Raw	Lower	Upper	Adult Lifespan	Age at Maturity	N_b Adjusted	N_e Adjusted
Study Area	2021-	0 to 1+	6.4	3.2	10.4	8.5	3	5.69	12.08
Study Area	2020-	2+	47.2	23.4	194	8.5	3	42.68	88.35

Table 5.3-13. Effective population size estimates for Dolly Varden corrected for overlapping generations using Waples et al. (2014) adjustment based on adult life span and age of first reproduction.

Population	Cohort	Age Class	N_b Raw	Lower	Upper	Adult Lifespan	Age at Maturity	N_b Adjusted	N_e Adjusted
Big Beaver	2020	1	23.00	16.90	32.7	8.5	2.5	20.72	43.13
Lightning	2021	0	14.50	10.00	22.1	8.5	2.5	13.02	27.24
Lightning	2020	1	24.70	18.00	35.4	8.5	2.5	22.26	46.31
Thunder	2021	0	33.70	26.6	44.10	8.5	2.5	30.42	63.13
Thunder	2020	1	27.90	17.3	56.6	8.5	2.5	25.16	52.29

Adult Lifespan and Age of Maturity taken from Jonsson et al. (1984). Only collections with greater than 20 samples were analyzed.

6.0 DISCUSSION AND FINDINGS

This study is complete and has met the goals and objectives stated in Section 2.0

6.1 Year 1

6.1.1 Summary of Completed Objectives

This section provides a summary of completed objectives for Year 1.

- City Light convened an Expert Panel. Expert Panel members are identified in Section 5.1 of this study report. The Expert Panel reviewed the Year 1 Existing Genetics Data Review Technical Memorandum (Attachment A) and provided guidance in the development of Year 2 study activities.
- The study team reviewed, compiled, and summarized genetics data collected in the Project reservoirs by multiple researchers. Specifically, the study team contacted the WDFW fish genetics laboratory and the USFWS Abernathy Fish Technology Center via email to request all genetic data and metadata. WDFW provided 2,697 *O. mykiss* genotypes for 15 microsatellite loci that appeared to have also been analyzed by Pflug et al. (2013). Ambiguity exists because individual identification for each genotype was not provided in the Pflug et al. (2013) report. WDFW and USFWS provided 898 char genotypes for 16 microsatellites that appeared to have been analyzed by Smith (2010).
- Limited information (metadata) was provided on how the samples were collected or what hypotheses were being tested by the existing data. Due to this ambiguity, the study team's consolidation efforts focused on reducing violation of statistical assumptions that are common to the analysis of microsatellite data in general. Specifically, efforts attempted to increase biologically meaningful signals within the data by reducing noise associated with (1) possible hybridization; (2) small sample sizes; (3) missing and erroneous data; and (4) violation of HWE and linkage equilibrium.
- The study team created a single, standardized data file for each species that compiles genotypes from existing studies. The genotypes were compiled into GENEPOP files (Raymond and Rousset 1995) that are available upon request.
- The study team used the standardized GENEPOP files to evaluate baseline genetic metrics for the three Project taxa for discussion purposes with the Expert Panel. Summaries of the review, compilation, and analysis for each taxon are provided in Sections 5.2.1 and 5.2.2 of this study report, which is content previously included in the ISR and Technical Memo to Expert Panel (Attachment D). (See the ISR for additional details on the existing information for *O. mykiss* and *Salvelinus* contained in Year 1 analyses and data reviews.)
- The study team calculated within- and among-population summary statistics using consistent methods for Bull Trout and Rainbow Trout.
- City Light estimated the power of genetic markers currently in use to identify relationships (e.g., parent-offspring pairs, full-sibling-unrelated pairs). For *O. mykiss* sampled in Roland Creek, a tributary within the study area, the FNR for identifying related individuals was 0.392. For Bull Trout sampled in Big Beaver, Ruby, and Stetattle creeks, the FNR estimated for

collections were 0.857, 0.868, and 0.95, respectively, meaning using existing microsatellite data for pedigree analysis is expected to result in more false relationship assignments than true assignments.

6.2 Year 2

6.2.1 Summary of Completed Objectives

This section provides a summary of completed objectives for Year 2.

- Sample collections were greatly expanded and coordinated during Year 2. During the 2022 field season, a total of 764 tissue samples were collected from *Oncorhynchus* and 342 from *Salvelinus*. In addition, 917 *Oncorhynchus* samples were provided by the USGS, 32 Bull Trout samples were provided by City Light, 180 microsatellite genotypes for Dolly Varden were provided by WDFW, and 876 *Oncorhynchus* GT-seq SNP genotypes were provided by WDFW. GT-seq SNP genotypes were not available for Bull Trout in the lower basin at the time this report was completed.
- Estimates of genetic diversity within and among Project reservoirs were completed.
- Additional data was gathered to estimate N_b and N_e for Bull Trout, Rainbow Trout and Dolly Varden. Estimates varied depending on how samples were grouped (e.g., by cohort, tributary, or inferred cluster) and whether estimates were corrected for bias associated with overlapping generations. For Dolly Varden, generational N_e within tributaries estimated from mixed cohort samples and uncorrected for bias was 23.40 (harmonic mean across Big Beaver, Colonial, Canyon, Granite, Lightning, and Ruby creeks). Bias corrected generational N_e was 42.81 (harmonic mean across Lightning, Big Beaver, and Thunder creeks, cohorts 2020 and 2021). For Bull Trout, samples from individual tributaries were too small to estimate N_b or N_e. However, considering the study area as a single collection, uncorrected, mixed cohort N_e=31.40 (95 percent CI 17.50, 68.80). Grouping samples into two cohort-based collections [age-0 to 1+] and [age 2+] produced N_e estimates of 12.08 and 88.35, respectively. Interpretation of N_e within metapopulations (i.e., with subpopulations exchanging gene flow) is not always straightforward and can be larger, smaller, or intermediate to the sum of their constituent subpopulations. Nevertheless, metapopulations also contain larger diversity as a whole relative to any individual subpopulations. To provide context, Ardren et al. (2011) estimated that 75 percent of Bull Trout populations in the U.S. are characterized by N_e<50, yet downstream of the Project area in the Sauk River, the lower 95 percent interval was approximately 200.
- Significant data needed to estimate N_e during the ILP study period was collected during the 2022 field season. Specifically, ages were estimated from scale analysis that were then combined with fork lengths recorded at time of capture. These data were used to generate an age-length key that could be applied to samples from the study area to estimate age given a fish's size. This key was applied to 2022 field collections by CFS, to partition genotype data into single-age cohorts required for a mathematical model used to estimate effective population size. This model can presumably be used and calibrated to serve future analysis efforts.

6.2.2 Discussion of Rainbow Trout (*Oncorhynchus mykiss*)

- Thirteen percent of randomly collected *Oncorhynchus* in the Project vicinity were hybrids with Cutthroat Trout, given resolution provided by three taxon-diagnostic loci. Meaning, Cutthroat and Rainbow Trout are reproducing together in the study area. Higher proportions of hybridized individuals were observed within collections from tributaries in the southern part of Ross Lake (e.g., Big Beaver).
- Three study area populations were observed to be highly genetically distinct: Little Beaver Creek, Three Fools Creek, and Pyramid Creek. Passage barriers could be reenforcing genetic distinctiveness of these populations. Little Beaver Creek collections occurred above a partial passage barrier. A partial barrier exists at the mouth of Lightning Creek, of which Three Fools is a tributary. There is a complete passage barrier present on Three Fools Creek, but collections occurred below this barrier. Meaning, individuals from Three Fools Creek had access to Lightning Creek, and some fish that were assigned to Three Fools Creek were recovered downstream in Lightning Creek collection. An analysis using the STRUCTURE program confirmed Lighting Creek individuals contained Three Fools ancestry (data not shown). Nevertheless, Three Fools Creek was distinct from adjacent Lightning Creek (and all other study area collections). Pyramid Creek collection occurred above a complete upstream passage barrier.
- In Ross Lake, Little Beaver Creek and Three Fools Creek populations are distinct and probably demographically independent, given their high degree of genetic differentiation (sensu Lowe and Allendorf 2010).
- Reproductive connection with Three Fools Creek is likely the source of Lightning Creek distinctiveness (STRUCTURE analysis, data not shown).
- There is a complete upstream passage barrier at the mouth of Pyramid Creek. Pyramid Creek fish were genetically aberrant in that this population appears to be derived from the coastal *O. mykiss* lineage (*O. m. irideus* subspecies). The coastal lineage is distributed widely along the Western U.S. and is the lineage of lower Skagit River *O. mykiss* (below Gorge Dam).
- The two genetic clusters widely distributed in the Project vicinity area (labelled cluster 3 and cluster 5) were stable across multiple years of collections (collections were combined across years for a location). As reproductive connection can homogenize genetic diversity within a short time (i.e., years), these clusters must be persisting through (non-random) assortative mating with respect to cluster identity. The two genetic clusters distributed throughout the Project vicinity area (labelled cluster 3 and cluster 5) were differentiated from each other, even for collections from the same location. For example, Granite Creek-3 and Canyon Creek-3 had a smaller distance (F_{ST} =0.005) between them than either Granite Creek-3 to Granite Creek-5 (distance=0.030) or Canyon-3 to Canyon-5 (distance=0.035). This pattern held for comparisons among lakes, with smaller distances (F_{ST}) observed between cluster-3 populations across Ross, Diablo, and Gorge lakes (Ruby, Canyon, Granite, Colonial, Thunder, Stetattle Creeks) than comparisons between cluster 3 and 5 from within the same tributary.
- Study results indicated that the current classification that there is a single population in the study area was not accurate.
- Categorizing the tributaries based on genetic distance (e.g., F_{ST}~0.02) is challenging given the presence of two genetic clusters (labeled 3 and 5). Considering just genetic cluster 3, the study

team suggests that these collections exhibit enough similarity to be treated as a single management unit, despite not having completely random mating. Yet, it seems unwise to ignore clusters 3 and 5, given the unknown qualities of their differences and dynamics of persistence.

- Given study area non-migratory (resident) *O. mykiss* are not protected under the state or federal Endangered Species Acts, and management decisions have considered the Project vicinity area a single population, the study team proposes initially classifying as four populations in Project vicinity: 1) Little Beaver Creek, 2) Three Fools Creek, 3) Lightning Creek, and 4) the remaining tributaries (excluding Pyramid Creek). At some future timepoint, management may have to account for reproductive dynamics between clusters 3 and 5 (at same location), which could necessitate altering the classification of populations to location by genetic cluster.
- Haplotype diversity was observed at chromosome 5 loci (OMY5), a location potentially associated with juvenile life-history, although said relationship is unknown in the Project vicinity area. Future evaluations of adaptive diversity could be conducted with data generated from this study and compared to other geographic regions. These study data could inform future deliberations regarding quantitative trait diversity present in the study area.
- Genetic distances were considerably higher for comparisons between Project vicinity area *O. mykiss* with collections representing populations from below Gorge Dam (upper Skagit River, Goodell Creek, lower Cascade River, Finney Creek, Marblemount Hatchery). Genetic distances (F_{ST}) were approximately an order of magnitude larger for tests between above Gorge Dam to populations below Gorge Dam. It is reasonable to conclude that Project vicinity area *O. mykiss* are distinct from Skagit River *O. mykiss*. Pairwise F_{ST} estimates among below-Project populations were less than 0.02, although the Marblemount Hatchery collection was more divergent, with F_{ST} = 3.5-5.1 when compared to other populations from below Gorge Dam.
- Project vicinity *O. mykiss* had lower genetic diversity than that observed for below-Project populations, although not remarkably so. The magnitude of genetic diversity present in Project vicinity area *O. mykiss* does not appear to be a concern, especially considering the reproductive connectivity observed within Project *O. mykiss* populations.
- A regional comparison was made for Project vicinity *O. mykiss* using a dataset consisting of 243 collections of *O. mykiss*. In total, 11,653 *O. mykiss* samples were included in this analysis. The results observed were similar to observations reported for the 2019 stranded *O. mykiss* analysis (Small et al. 2020). Project vicinity *O. mykiss* were intermediate (on first genetic principal component) with respect to coastal (*O. m. irideus*) and inland redband (*O. m. gairdneri*) ancestry. Additionally, the second genetic principal component represented genetic variation between Project vicinity *O. mykiss* and all other populations present in reference database. Project vicinity *O. mykiss* genetic characteristics appear unique compared to other populations within Washington State.
- The uniqueness of Project vicinity *O. mykiss* has implications for discussions regarding human-mediated passage.

6.2.2.1 Conclusions

Study area *O. mykiss* had lower diversity than that observed for populations from below Gorge Dam, but not remarkably so. The annual effective number of breeders (N_b) was estimated from

some study area locations where samples numbers were sufficient for 2022 age-1 cohort. Initial estimates of N_b were calculated from the widely distributed genetic cluster 3 (excluding Lightning Creek), Little Beaver Creek, Pyramid Creek, and Stetattle Creek genetic cluster 5. The N_b from the amalgamated genetic cluster 3 was 394.9 (95 percent CI 321.0-508.5). Little Beaver Creek (genetic cluster 1) age-1 cohort had an estimated N_b of 106.2 (95 percent CI 53.8-1144.1). Pyramid Creek (genetic cluster 2) age-1 cohort had an estimated N_b of 13.5 (95 percent CI 12.5-14.7). Stetattle Creek (genetic cluster 5) age-1 cohort had an estimated N_b of 28.7 (95 percent CI 26.2-31.5). Given reproductive connectivity, lower genetic diversity observed is not a concern.

Three study area populations were observed to be highly genetically distinct: Little Beaver Creek, Three Fools Creek, and Pyramid Creek. Two additional genetic clusters widely distributed throughout the study area were differentiated from each other, even for collections obtained from the same location. The study results indicated that the current classification that there is a single population in the study area was not accurate. The study team proposes initially classifying *O. mykiss* in the Project vicinity as: (1) Little Beaver Creek; (2) Three Fools Creek; (3) Lightning Creek; and (4) the remaining tributaries (excluding Pyramid Creek).

Genetic distances were an order of magnitude higher for comparisons between study area *O. mykiss* with collections representing populations from below Gorge Dam. It is reasonable to conclude that Project vicinity *O. mykiss* are distinct from Skagit River *O. mykiss*. Pyramid Creek fish were genetically aberrant in that this population appeared to be derived from the coastal *O. mykiss* lineage (i.e., from below Gorge Dam). Study area *O. mykiss* appear unique compared to other populations from Washington State.

6.2.3 Discussion of Native Char (*Salvelinus* spp.)

- SNP genotypes from the lower basin (downstream of Gorge Lake) were requested from WDFW and USFWS but were not received in time to make it into this study report. The study team received the lower basin SNP genotypes on January 31, 2023 and plans to analyze them early spring 2023. Results will be shared with the Expert Panel and LPs possibly as early as March or April 2023.
- Sample sizes in Bull Trout during year 2 were too small to analyze by tributary and so it is important to consider results as preliminary. During year 2, genetic diversity measured as the proportion of polymorphic loci within collections of Bull Trout in the study area was relatively low compared to Bull Trout immediately downstream in the mainstem Skagit River. Forty GT-Seq SNP loci were polymorphic in a single Bull Trout collected near Marblemount compared to 33 in a collection of 65 Bull Trout in the study area. Although a sample size of one fish is too small to make any statistical inferences, the result aligns with the Year 1 analysis of microsatellites and with a recent mtDNA study completed by USFWS (Smith). Smith (USFWS, unpublished) hypothesized that colonization by Bull Trout within the study area could have resulted from ancient capture of the Skagit River by the Fraser River during the last ice age. Such an event would be expected to result in reduced genetic variation of contemporary Bull Trout within the study area via a genetic bottleneck. Nevertheless, reference samples from the Fraser and downstream of the Project would be needed to substantiate this hypothesis. A single Bull Trout is insufficient to make robust conclusions about the genetic diversity of Project area Bull Trout relative to those downstream.

- Inference about the genetic diversity of Dolly Varden compared to populations across their range is unclear because reference samples of Dolly Varden from nearby populations were not available for comparison. Nevertheless, the heterozygosity at a subset of four microsatellites analyzed in collections of southern Dolly Varden from across their range was relatively similar: the range-wide H_S was 0.63, compared to 0.60 in the Project vicinity (Taylor et al. 2015). No statistical test was implemented to compare these values.
- Genetic differences among Bull Trout collections within the study area was apparent. Grouping Bull Trout into *a priori* collections reflecting current reservoir boundaries explained a significant proportion of genetic variation in the Project (AMOVA: F_{ST}=0.05; P less than 0.01) and collections did not deviate significantly from HWE (F_{IS}=-0.01; P=0.66). Pooling samples was necessary because sample sizes were too small to analyze by tributary. Nevertheless, unsupervised analysis of genetic structure using DAPC did not consistently group Bull Trout into collections based on current reservoir boundaries, suggesting contemporary reservoir boundaries do not explain all the genetic structure within the study area. Specifically, the most supported DAPC model placed Bull Trout into three genetic clusters that were somewhat well mixed between Ross and Gorge lakes, although Bull Trout from Ross tended to have higher loadings for the first linear discriminant relative to those from Gorge (Figure 5.3-22).
- Genetic divergence among Dolly Varden collections was also apparent. The study team observed a strong pattern of isolation-by-distance (F_{ST} ranged from 0.003 to 0.081), which has been observed in Dolly Varden from other watersheds (Melnik et al. 2019) and suggests that gene flow between proximate sites is more likely than distal sites. Isolation by distance is consistent with migration drift equilibrium, yet the Mantel test of the residuals was nonsignificant, suggesting the presence of equilibrium is ambiguous. A population at migration-drift equilibrium is a theoretical state where the opposing forces of migration and genetic drift have reached a balance. The observed degree of divergence could indicate demographic independence of some populations of Dolly Varden in the Project area (sensu Lowe and Allendorf 2010). DAPC suggested 12 genetic clusters was the most parsimonious model of genetic structure in the study area. However, the 12 genetic clusters showed little association with contemporary watershed boundaries, which was unexpected in the face of isolation-by-distance (i.e., because isolation by distance implies genetic structure associated with watershed boundaries). Unlike STRUCTURE, DAPC places individuals into clusters based on allelic state and does not assume HWE within collections. In cases where isolation by distance is present, clustering algorithms can produce misleading signals of genetic structure or grouping because the nature of these algorithms may not match the continuous isolation by distance structure, which may not be characterized by discrete populations (Perez et al. 2018).
- Estimates of N_e in native *Salvelinus* were small no matter how collections were grouped (e.g., by cohort or mixed) or corrected for bias. All point estimates were less than 100 and most were less than 50. Small N_e is common in Bull Trout due to their breeding ecology (i.e., typically few adults achieve relatively high reproductive success). Additionally, Bull Trout sample size to support this analysis was low. N_e and N_b estimates for Dolly Varden in the study area were also small. Like Bull Trout, all point estimates were less than 100 and most were less than 50.
- The 50/500 rule is a general rule-of-thumb in conservation science that states Ne should not be less than 50 in the short-term, and not less than 500 in the long-term. The short-term rule is based on well-documented decreases in fitness due to inbreeding when N_e falls below

approximately 50. The long-term rule is based on the loss of adaptive genetic variation that is important for potential local adaptation. In the context of a metapopulation (e.g., native trout and char in the study area), Laikre et al. (2016) recommended that long-term genetic viability should imply that the rate of inbreeding in the entire metapopulation (Ne_{Meta}), as well as in the separate subpopulations (N_{eX}), should be greater than 500 due to the risk of accumulation of inbreeding within subpopulations. Interpretation of N_e within metapopulations (i.e., with subpopulations exchanging gene flow such as in the study area) is not always straightforward and can be larger, smaller, or intermediate to the sum of their constituent subpopulations (Allendorf et al. 2013). Complicating inferences about the threat of inbreeding due to small effective size within the study area is that hybridization with Dolly Varden appears somewhat common.

- To estimate N_b, the length-at-age model was constructed at the genus level instead of the species level. One reason is because hybrid identification in the field can be unreliable, and the model was built during the field season due to the protracted reporting deadline. Additionally, only 30 Bull Trout were sampled, which is not a sufficient sample size to construct a reliable model at the species level. One disadvantage of using a mixed sample is that it could lead to statistical and inferential biases. A model built with individuals of mixed ancestry could bias the model and this bias would depend on specific characteristics of the different species within the genus. For example, there may be species-specific differences in growth rates or mortality patterns that would be averaged out when using a mixed sample. The main reason a mixed sample was used is because so few Bull Trout were captured, and due to the overlap in morphology between different species, it was unreliable to accurately differentiate between species in the field. Thus, care should be taken interpreting N_b results.

6.2.3.1 Conclusion

Inferences should be considered preliminary due to small sample sizes and absence of samples from the lower basin. The SNP analysis of Bull Trout populations in the study area showed genetic structure among them when grouped by reservoir (F_{ST}=0.05). The analysis also showed that a significant amount of the genetic structure among Bull Trout populations could be explained by contemporary reservoir boundaries. However, telemetry data has provided limited evidence that Bull Trout have dispersed downstream through the reservoirs, and unsupervised analysis of genetic structure demonstrated that reservoirs do not account for all the structure. Year 1 microsatellite data clearly showed that Bull Trout from the Project area are highly genetically distinct from those downstream of Gorge Lake, as evidenced by the exceptionally high F_{ST} values (F_{ST} ranged from 0.27 to 0.41). The new GT-seq SNP data provided in this report showed that a single Bull Trout from Year 2 had the highest proportion on polymorphic SNPs in the entire dataset. More genotypes from downstream would be needed to substantiate this as a consistent pattern using the GT-seq SNP panel. Bull trout in the Project area have lower genetic diversity compared to downstream populations, which is supported by Year 1 microsatellites and an mtDNA study completed by M. Smith (USFWS unpublished). With respect to year 2, the single Bull Trout sampled from downstream of the Project area was a genetically distinct outlier, accounting for nearly all the inertia of the first PC. Again, comparison to more genotypes from downstream of the project area would be needed to substantiate this pattern as consistent at the GT-seq SNP panel. Hybridization between native *Salvelinus* species was found to be more common than hybridization with invasive Brook Trout. Microsatellite analysis of Dolly Varden revealed a strong pattern of isolation-by-distance. Estimates of N_e in native *Salvelinus* were small, with most estimates less than 50. Small

N_e is common in Bull Trout due to their breeding ecology, and it is recommended to conserve interconnected subpopulations at least large enough to meet 1000 spawners and/or the 50/500 rule. Metapopulations composed of multiple small populations, including subpopulations from Canada, could harbor more genetic diversity than expected due to exchange. The sample size of 65 individuals may not be sufficient to accurately represent the entire population and additional samples from the Project area, downstream, and Canada would provide a more accurate representation of the genetic diversity, genetic structure, and effective population size of the Bull Trout population.

6.3 Status of June 9, 2021 Notice

The June 9, 2021 Notice noted five items of discussion related to the implementation of the Reservoir Fish Genetics Study. The status of each is summarized in Table 6.3-1.

Table 6.3-1. Status of Stranding and Trapping Assessment modifications identified in the June 9, 2021 Notice.

Study Modifications Identified in the June 9, 2021 Notice: As Written	Status
SCL will modify study plan to collect juvenile fish at spawning grounds for genetics baseline as part of field sampling program in Year 2. Action item: SCL to modify study plan and circulate to LPs after FERC's issuance of the study plan determination.	Collection of juveniles (young-of-year) on spawning and nursery grounds was the approach included in the 2022 Proposed Year 2 Sampling Plan (Attachment B), which was shared with LPs and the Expert Panel in April 2022. This was completed in Year 2 of the Reservoir Fish Genetics Study.
SCL will modify study plan to expand sample collection/coordination of existing samples and activities and analysis out of basin and above/below dams.	Regional (within and outside of Skagit basin) and above and below-dam analyses within the Project area were completed in Year 2 of the Reservoir Fish Genetics Study for *Oncorhynchus*. Comparisons to collections from downstream of the Project area were not included in this study report for *Salvelinus* as the data were not provided in time for reporting. This analysis will be conducted in Spring 2023 and shared with the Expert Panel.
SCL will clarify study plan to explain the role of the expert panel. The LPs and SCL agree that: 1) the expert panel will serve in an advisory role, and 2) the expert panel will include experts from fields other than genetics.	City Light clarified to LPs and Expert Panel members that the role of the Expert Panel is advisory. A variety of experts including resource agency specialists and experts from academia with backgrounds in genetics and/or ecology were included on the Expert Panel.
SCL will modify FA-06 to provide that SCL will seek input from LPs and advice from an expert panel on whether and how genetics information or other monitoring methods can be used to inform future evaluation of reservoir fish abundance, habitat use, and migration timing.	City Light has sought the input of the Expert Panel on the implementation of this study. City Light expects that this study will provide useful information to inform future management decisions.
This issue [that Project operations and resource effects section is missing from study plan] will be addressed in the Draft License Application.	This issue was addressed in the Draft License Application and will be addressed in the Final License Application.

7.0 VARIANCES FROM PROPOSED STUDY PLAN AND PROPOSED MODIFICATIONS

The Reservoir Fish Genetics Study was not a FERC-required study and was implemented voluntarily by City Light.

8.0 REFERENCES

Allendorf, F. W., G. Luikart, and S. N. Aitken. 2013. Conservation and the genetics of populations. Wiley-Blackwell.

Anderson, S. 2019. CKMRsim package, efficient sampling for close-kin mark recapture power estimation. [Online] URL: https://github.com/eriqande/CKMRsim. September 2021.

Angers, B. and L. Bernachez. 1996. Usefulness of heterologous microsatellites obtained from brook charr, *Salvelinus fontinalis* Mitchill, in other *Salvelinus* species. Molecular Ecology 5:317–319.

Ardren, W.R., DeHaan, P.W., Smith, C.T., Taylor, E.B., Leary, R., Kozfkay, C.C., Godfrey, L., Diggs, M., Fredenberg, W., Chan, J. and Kilpatrick, C.W., 2011. Genetic structure, evolutionary history, and conservation units of bull trout in the coterminous United States. Transactions of the American Fisheries Society, 140(2), pp.506-525.

Banks, M. A., M. S. Blouin, V. K. Baldwin, H. A. Rashbrook, H. A. Fitzgerald, S. M. Blankenship, and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in Chinook Salmon (*Oncorhynchus tschawytscha*). The Journal of Heredity 2.

Beauchamp, D. (in development). Factors limiting native salmonids above Skagit River dams study (Food Web Study). U.S. Geological Survey (USGS).

Bohling, J., J. Von Bargen, M. Piteo, A. Louden, M. Small, T. A. Delomas, and R. Kovach. 2021. Developing a Standardized Single Nucleotide Polymorphism Panel for Rangewide Genetic Monitoring of Bull Trout. North American Journal of Fisheries Management 41(6):1920–1931.

Buchholz, W. G., S. J. Miller, and W. J. Spearman. 2001. Isolation and characterization of chum salmon microsatellite loci and use across species. Animal Genetics 32.

Cairney, M., J. B. Taggart, and B. Hoyheim. 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. Molecular Ecology 9:2155-2234.

Campbell, N. R., S. A. Harmon, and S. R. Narum. 2014. Genotyping-in-Thousands by sequencing (GTseq): A cost effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources 15(4):855–867.

Copeland, T., K. Hernandez, M. T. Davison, and K. Wright. 2018. Validation of spawn checks and saltwater age assignments based on scales from known repeat-spawning steelhead. North American Journal of Fisheries Management 38(5):1050–1058.

Crane, P. A., C. J. Lewis, E. J. Kretschmer, S. J. Miller, W. J. Spearman, A. L. DeCicco, M. J. Lisac, and J. K. Wenburg. 2004. Characterization and inheritance of seven microsatellite loci from Dolly Varden, *Salvelinus malma*, and cross-species amplification in Arctic char, S. alpinus. Conservation Genetics 5:737-741.

DeHaan P. W. and W. R. Ardren. 2005. Characterization of 20 highly variable tetranucleotide microsatellite loci for bull trout (*Salvelinus confluentus*) and cross amplification in other Salvelinus species. Mol. Ecol. Notes 5: 582-585.

DeHaan, P., B. J. Adams, and M. C. Nelson. 2014. Fine-scale population structure analysis and genetic population assignments of Wenatchee River sub-basin bull trout. USFWS, editor, Longview, WA.

Earl, D. A., and B. M. VonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation genetics resources 4:359–361.

Excoffier, L., and M. Slatkin. 1998. Incorporating genotypes of relatives into a test of linkage disequilibrium. Am. J. Hum. Genet. 171-180.

Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Fisher, J. 2020. Seattle City Light correspondence.

Fisher, J. 2022. Seattle City Light correspondence.

Franklin, I. R. 1980. Evolutionary change in small populations. Pages 135–150 Conservation Biology An Evolutionary-Ecological Perspective. Soule ME, Wilcox BA. Sinauer, Sunderland, Massachusetts.

Garant, D., J. J. Dodson, and L. Bernatchez. 2000. Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar L.*). Molecular Ecology 9(5):615–628.

Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of Heredity 86:485-486.

Goudet, J., M. Raymond, T. de Meeüs, and F. Rousset. 1996. Testing differentiation in diploid populations. Genetics 144: 1933-1940.

Guo, S. and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hale, M. L., T. M. Burg, and T. E. Steeves. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population Is enough to accurately estimate allele frequencies. PLoS ONE 7:1-10.

Hemmingsen, A. R., Gunckel, S. L., Sankovich, P. M., & Howell, P. J. (2001). Bull trout life history, genetics, habitat needs, and limiting factors in central and northeast Oregon. Annual Report (1997): 95-54.

Hess, J.E., S.A. Harmon, D.J. Hasselman, R.L. Horn, A.R. LaGrange, A.P. Matala, and S.R. Narum. 2018. Genetic Assessment Of Columbia River Stocks, 1/1/2018 - 12/31/2018. Annual Report for Bonneville Power Administration, 2008-907-00.

Hernandez, K., T. Copeland, and K. Wright. 2014. Quantitative Assessment of Scale Resorption in Migrating and Spawning Steelhead of the Snake River Basin. Transactions of the American Fisheries Society 143(6):1562–1568.

Hutchison, D. W., and A. R. Templeton. 1999. Correlation of Pairwise Genetic and Geographic Distance Measures: Inferring the Relative Influences of Gene Flow and Drift on the Distribution of Genetic Variability. Evolution 53(6):1898–1914.

Isermann, D. A., and C. T. Knight. 2005. A Computer Program for Age–Length Keys Incorporating Age Assignment to Individual Fish. North American Journal of Fisheries Management 25(3):1153–1160.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21):3070–3071.

Jombart,T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal 732 components: a new method for the analysis of genetically structured populations. BMC 733 Genet. 11:94. BioMed Central Ltd.

Jombart, T. 2011. *adegenet*: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10:551-555..

Jonsson, B., Hindar, K. N. T. G., and Northcote, T. G. (1984). Optimal age at sexual maturity of sympatric and experimentally allopatric cutthroat trout and Dolly Varden charr. Oecologia, 61(3): 319-325.

Laikre, L., F. Olsson, E. Jansson, O. Hössjer, and N. Ryman. 2016. Metapopulation effective size and conservation genetic goals for the Fennoscandian wolf (*Canis lupus*) population. Heredity 117(4):279–289.

Love, D. C. 2016. Manual for aging steelhead trout scales based on scale sampling from Sitkoh Creek, Big Ratz Creek, and other Southeast Alaska streams, 2003-2011. Alaska Department of Fish and Game, Division of Sport Fish, Research and Technical Services.

Lowe, W. H., and F. W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19(15):3038–3051.McConnell, S., L. Hamilton, D. Morris, D. Cook, D. Paquet, P. Bentzen, and J. Wright. 1995. Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. Aquaculture 137:19-30.

Melnik, N. O., G. N. Markevich, E. B. Taylor, A. V. Loktyushkin, and E. V. Esin. 2020. Evidence for divergence between sympatric stone charr and Dolly Varden along unique environmental gradients in Kamchatka. Journal of Zoological Systematics and Evolutionary Research 58(4):1135–1150.

McConnell, S. K., O'Reilly, P., Hamilton, L., Wright, J. M., & Bentzen, P. 1995. Polymorphic microsatellite loci from Atlantic salmon (Salmo salar): genetic differentiation of North American and European populations. Canadian Journal of Fisheries and Aquatic Sciences, 52(9), 1863-1872.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY, USA.

Nelson, R. J., T. D. Beacham, and M. P. Small. 1998. Microsatellite analysis of the population structure of a Vancouver Island sockeye salmon (*Oncorhynchus nerka*) stock complex using nondenaturing gel electrophoresis. Molecular Marine Biology and Biotechnology 7:312-319.

Ogle, D. H., J. C. Doll, P. Wheeler, and A. Dinno. 2022. FSA: Fisheries Stock Analysis. R package version 375 0.9. 3.

Olsen, J. B., P. Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7:1087-1089.

Ostberg, C. O., and Chase, D. M. 2022. Ontogeny of eDNA shedding during early development in Chinook Salmon (Oncorhynchus tshawytscha). Environmental DNA, 4(2), 339-348.

Pearse, D. E., M. R. Miller, A. Abadía-Cardoso, and J. C. Garza. 2014. Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. Proceedings of the Royal Society B: Biological Sciences 281(1783):20140012.

Perez, M.F., Franco, F.F., Bombonato, J.R., Bonatelli, I.A., Khan, G., Romeiro-Brito, M., Fegies, A.C., Ribeiro, P.M., Silva, G.A. and Moraes, E.M., 2018. Assessing population structure in the face of isolation by distance: are we neglecting the problem?. Diversity and Distributions, 24(12), pp.1883-1889.

Pflug, D. E., E. J. Connor, B. Hayman, T. W. Kassler, K. Warheit, B. McMillan, and E. Beamer. 2013. Ecological, genetic, and productivity consequences of interactions between hatchery- and natural-origin steelhead of the Skagit watershed. Skagit River System Cooperative.

Queller, D.C. and K.F. Goodnight. 1989. Estimating relatedness using genetic markers. Evolution 43(2):258–275.

Raymond M. and F. Rousset. 1995. An exact test for population differentiation. Evolution 49:1280-1283

Reeves, P. A., C. L. Bowker, C. E. Fettig, L. R. Tembrock, and C. M. Richards. 2016. Effect of error and missing data on population structure inference using microsatellite data. BioRxiv. 080630.

Rexroad, C. E. III, R. L. Coleman, A. M. Martin, W. K. Hershberger, and J. Killefer. 2001. Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Animal Genetics 32:317-319.

Reynolds, J. B., A. L. Kolz, A. V. Zale, D. L. Parrish, and T. M. Sutton. 2012. Electrofishing. Pages 305–361.

Rice, W.R. 1989. Analyzing Tables of Statistical Tests. Evolution, 43: 223-225.

Rousset, F. 2020. Genepop Manual. Version 4.7.5.

Rousset, F. 2020. Population Genetic Data Analysis Using Genepop, R package, version 1.1.2.

Ryman, N., and S. Palm. 2006. POWSIM: A computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6:600-602.

Ryman, N., L. Laikre, and O. Hössjer. 2019. Do estimates of contemporary effective population size tell us what we want to know? Molecular Ecology 28(8):1904–1918

Scribner, K. T., J. R. Gust, and R. L. Fields. 1996. Isolation and characterization of novel microsatellite loci: cross species amplification and population genetic applications. Canadian Journal of Fisheries & Aquatic Sciences 53.

Seamons, T. 2020. Washington Department of Fish and Wildlife correspondence.

Seattle City Light (City Light). 2021. Revised Study Plan (RSP) for the Skagit River Hydroelectric Project, FERC Project No. 553. April 2021.

_____. 2023a. FA-03 Reservoir Fish Stranding and Trapping Risk Assessment, Updated Report for the Skagit River Hydroelectric Project, FERC Project No. 553. Prepared by Four Peaks Environmental Science & Data Solutions and Blue Coast Engineering. March 2023.

Sinnwell, J. P., D. J. Schaid, and M. J. P. Sinnwell. 2022. Package 'haplo. stats.'

Small, M. P., E. Lowery, and M. Kissler. 2020, March 12. Genetic analysis of char from Gorge Lake, Upper Skagit River, Washington. Washington Dept. of Fish and Wildlife Final Report.

Small, M.P., E. Martinez, and S. Peterson. 2013. Genetic analysis of native char collected in Diablo Lake, Ross Lake, Gorge Reservoir, Skagit mainstem, and Skagit Bay. WDFW Molecular Genetics Lab, Olympia WA.

Small, M.P., S. Bell, and C. Bowman. 2016. Genetic analysis of native char collected in Diablo Lake, Gorge Reservoir, and Ross Lake in the Skagit River basin. Washington Department of Fish and Wildlife Molecular Genetics Lab, Olympia, Washington.

Smith, C. T., B. F. Koop, and R. J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7:1614-1616.

Smith, M. J. 2010. Population structure and genetic assignment of bull trout (*Salvelinus confluentus*) in the Skagit River Basin. School of Aquatic and Fishery Sciences University of Washington.

Smith, M. 2021. Personal communication between Scott Blankenship, Science Director Cramer Fish Sciences and Matt Smith, Fish Geneticist USFWS. June 14, 2021.

Spies, I. B., D. J. Brasier, P. T. L. O'Reilly, T. R. Seamons, and P. Bentzen. 2005. Development and characterization of novel tetra-, tri-, and dinucleotide microsatellite markers in rainbow trout (*Oncorhynchus mykiss*). Molecular Ecology Notes 5:278-281.

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

Taylor, E.B., M.D. Stamford, and J.S. Baxter. 2003. Population subdivision in westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) at the northern periphery of its range: evolutionary inferences and conservation implications. Molecular Ecology, 12(10), 2609–2622.

Taylor, E.B., and S.L. May-McNally. 2015. Genetic analysis of Dolly Varden (*Salvelinus malma*) across its North American range: evidence for a contact zone in southcentral Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 72(7), 1048–1057.

Waples, R., and E. C. Anderson. 2017. Purging putative siblings from population genetic data sets: a cautionary view. Molecular Ecology 26:1211-1224.

Waples, R., and O. E. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419–1439.

Waples, R. S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics (7):167–184.

Waples, R. S., T. Antao, and G. Luikart. 2014. Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. Genetics 197(2):769–780.

Waples, R. S., and C. Do. 2008. ldne: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8(4):753–756.

Washington Department of Fish and Wildlife. 2021. *Salvelinus* genotype microsatellite data. Unpublished.

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

Weir, B.S. and J. Goudet. 2017. A Unified Characterization of Population Structure and Relatedness. Genetics 206(4):2085–2103.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-Statistics for the Analysis of Population-Structure. Evolution 38:1358–1370.

Whiteley, A. R., J. A. Coombs, M. Hudy, Z. Robinson, K. H. Nislow, and B. H. Letcher. 2012. Sampling strategies for estimating brook trout effective population size. Conservation Genetics 13(3):625–637.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics. 15:323–354.

Yang, R.-C. 1998. Estimating Hierarchical F-Statistics. Evolution 52(4):950–956.

RESERVOIR FISH GENETICS STUDY

ATTACHMENT A

YEAR 1 EXISTING GENETICS DATA REVIEW TECHNICAL MEMORANDUM

Technical Memorandum

Date:	Tuesday, January 04, 2022
Project:	Skagit River Hydroelectric Project
To:	Seattle City Light
From:	Scott Blankenship and Dan Bingham, Cramer Fish Sciences
Subject:	FA-06 Reservoir Native Fish Genetics Baseline Study – Existing Genetics Data Review

1.0 INTRODUCTION

1.1 Introduction

The Skagit River Hydroelectric Project is owned and operated by Seattle City Light (City Light) and is undergoing Federal Energy Regulatory Commission (FERC) relicensing with its current operating license expiring in 2025. The FA-06 Reservoir Native Fish Genetics Baseline Study (Reservoir Fish Genetics Study) was not required by FERC in its July 16, 2021 Study Plan Determination; however, City Light is implementing the Reservoir Fish Genetics Study as proposed in the Revised Study Plan (RSP; City Light 2021) with the agreed upon modifications described in the "Notice of Certain Agreements on Study Plans for the Skagit Relicensing" (June 9, 2021 Notice)[1]. The Year 1 objectives outlined in the Reservoir Fish Genetics Study aim to use genetic data produced by previous studies to better understand the types of inferences that can be drawn about the genetic diversity and population structure of native fish and to identify possible data gaps that might prevent satisfactory answers to questions submitted by licensing participants (LP) to City Light. Year 2 will use inferences gleaned from the Year 1 efforts (i.e., information provided in this memo) to fill any data gaps and provide answers to any outstanding questions pursued.

The microsatellite genotypes analyzed in this memo were previously identified in the FA-06 Reservoir Fish Genetics Study Plan as potentially providing useful baseline genetic data for native salmonids (City Light 2021). Since the early 2000s, microsatellite data were collected by multiple researchers from native salmonids sampled in the Project reservoirs, their tributaries, and from outside the Project vicinity but within the Skagit and nearby basins. City Light worked with primary researchers that produced these datasets (U.S. Fish and Wildlife Service [USFWS] and Washington Department of Fish and Wildlife [WDFW]) and with CFS to obtain and compile those existing native salmonid genetics data and to unify them for a common analysis of the baseline genetics metrics identified in subsequent sections of the RSP. This memo therefore reflects a *post hoc* analysis of existing microsatellite genotypes from Rainbow Trout (*Oncorhynchus mykiss*) and Bull Trout (*Salvelinus confluentus*) collected from inside and outside of the Project reservoirs

[1] Referred to by FERC in its July 16, 2021 Study Plan Determination as the "updated RSP."

(Figures 1 and 2) and is intended to establish a basis of existing information to support the Reservoir Fish Genetics Study.

Data requests were made by City Light to obtain the previously identified and pertinent microsatellite genotypes listed in Table 2.5-1 of the RSP. The genotypes for Rainbow Trout were originally produced and analyzed by Pflug et al (2013). Most genotypes for Bull Trout were produced and analyzed by Smith (2010), but genotypes from Small et al. (2013), (2016) and (2020) were also included. Methods described in this memo were not chosen to test hypotheses of the original studies, to provide a peer-review, nor to implement meta-analysis of their results. Rather, methods were chosen to describe the existing data in the context of the Reservoir Fish Genetics Study objectives. Due to inherent risks associated with evaluating existing genetic information for *post hoc* scientific investigations in general (i.e., inference based on analyses outside of the intended scope of the initial study design), caution is warranted during interpretation of results presented in this memo.

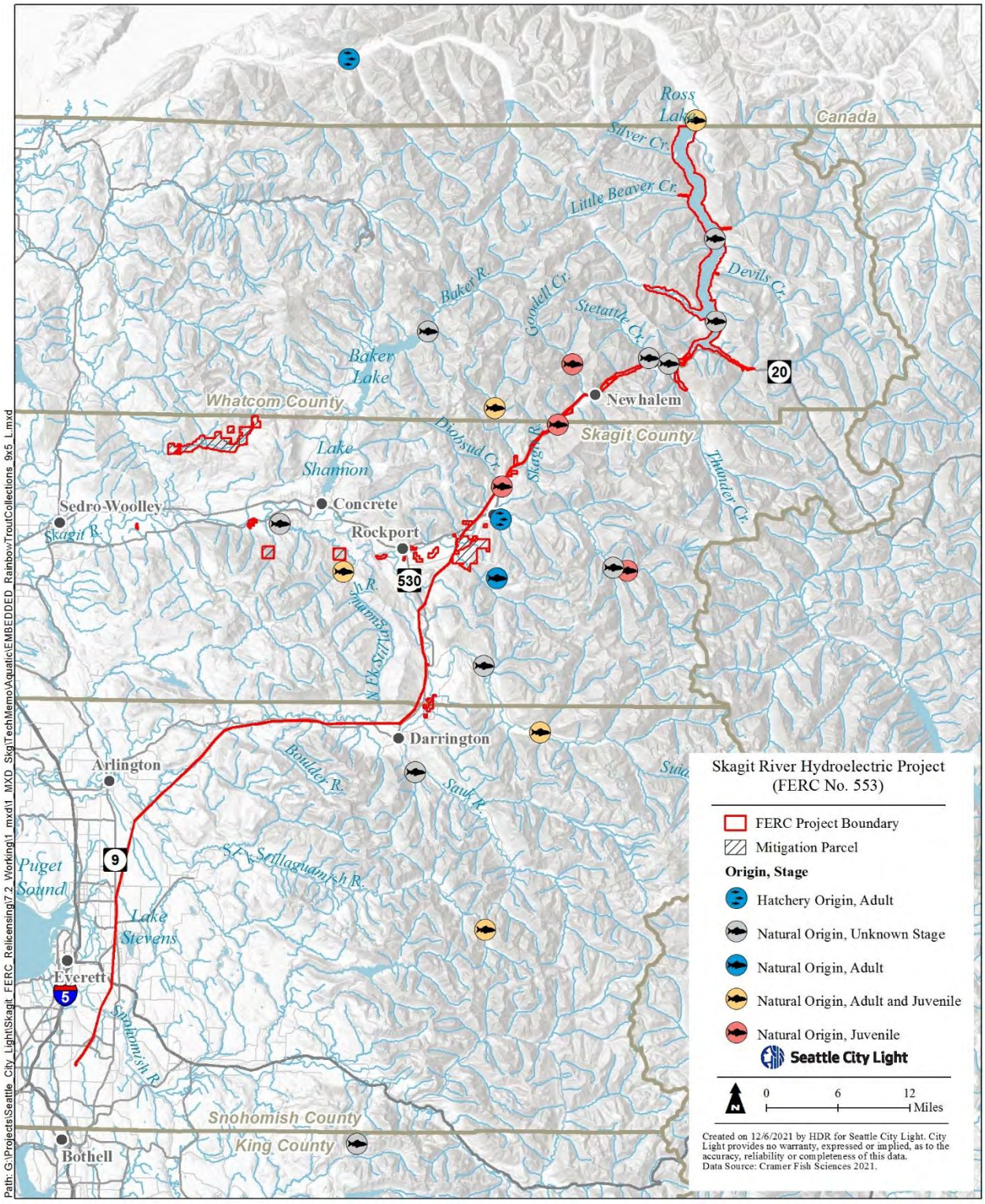


Figure 1. Proposed study area and collections evaluated for Rainbow Trout.

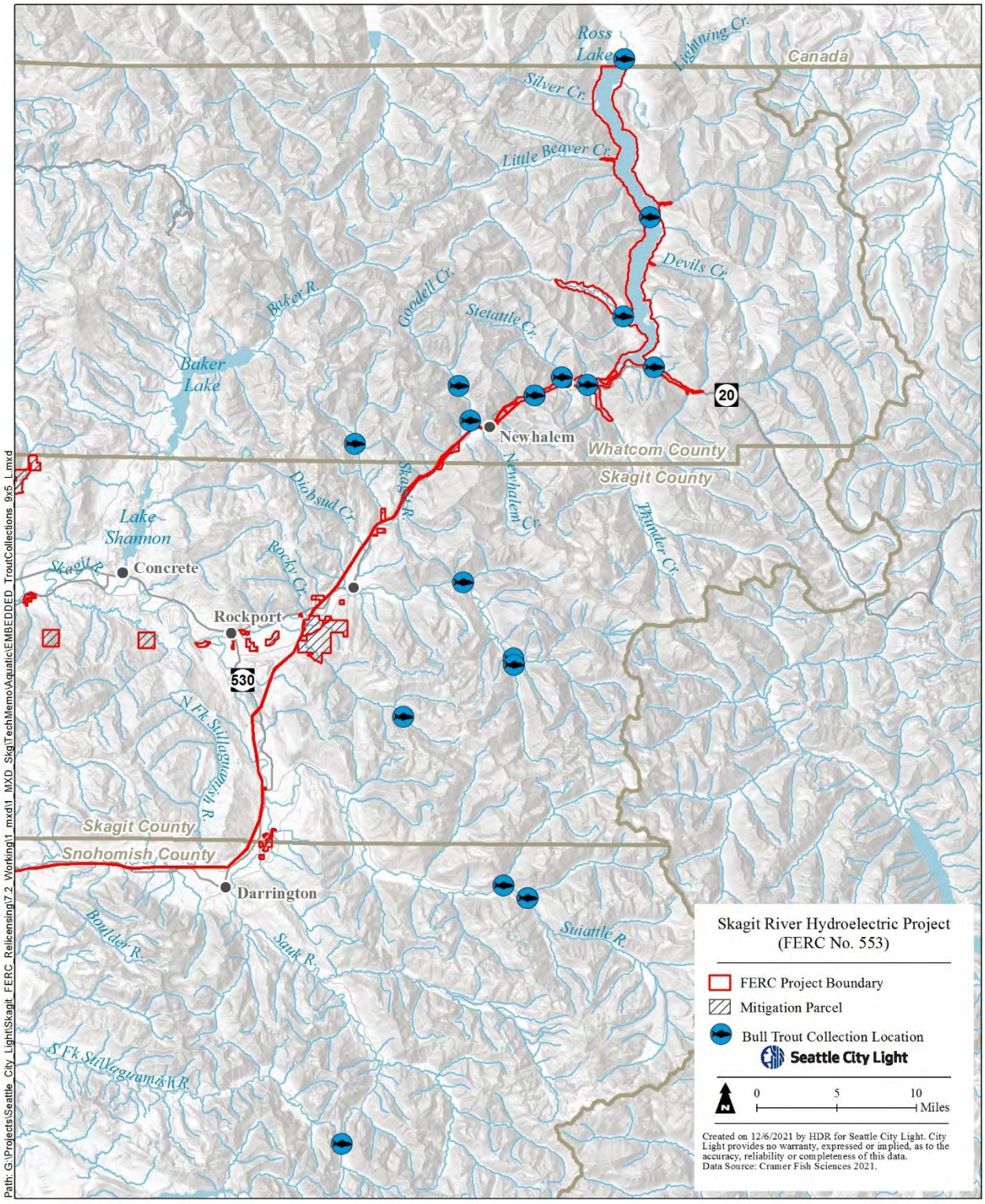


Figure 2. **Proposed study area and collections evaluated for Bull Trout.**

2.0 METHODS

2.1 What preexisting genetic data are available?

On June 6, 2021, Cramer Fish Sciences (CFS) emailed Todd Seamons of the WDFW genetics laboratory (Olympia, WA) and Matt Smith of the USFWS Abernathy Fish Technology Center (Longview, WA), requesting the microsatellite genotypes and metadata analyzed in Pflug et al. (2013), Smith et al. (2010), Small et al. (2013), (2016) and (2020). On September 9, 2021, CFS sent an additional request for any geospatial data that could aid in identifying the specific locations that tissue samples were collected. Todd Seamons and Matt Smith each forwarded the requested genotypes (See Section 3 of this memo for details) but indicated geospatial data were unavailable, and thus, precise reaches of rivers, streams, or positions in a lake are not known beyond descriptions in the original reports.

2.2 What types of inference can be drawn?

Understanding the designs of the original studies is important because any new conclusions are naturally limited by the initial scopes of inference (e.g., sampling designs, genetic marker choices, etc.). Briefly, the study design of Pflug et al. (2013) was observational in nature (i.e., not experimental). Sampling occurred between 2008 and 2010 and appeared to be loosely stratified by life history (i.e., anadromous versus resident), life stage (juvenile versus adult), degree of isolation (upstream versus downstream of a migration barrier), and by origin (hatchery- versus natural-origin), but allocation of sampling effort within strata was not defined *a priori*. The extent to which sampling was random, opportunistic, or targeted was unclear, but the report stated that for each collection, attempts were made to obtain 100 adult steelhead, 100 juvenile steelhead from anadromous zones, and 100 resident rainbow trout from above barriers. The geographic extent of sampling was broad across the Skagit River basin, including multiple collections from above and below the Project Boundary and within the Project reservoirs (Figures 1 and 2). Thus, some collections may characterize subpopulations while others may be mixtures. Samples from natural populations in the Sauk and Fraser rivers were also examined and so were samples from hatchery populations commonly used to supplement natural populations of Rainbow Trout in the Skagit River. Fifteen microsatellite markers were analyzed to "*…provide information about basic genetic characteristics of natural and hatchery origin steelhead populations and resident [O.] mykiss populations.*" Nine additional objectives that can be broadly categorized as descriptions of genetic population structure were also listed but no testable hypotheses were defined. The statistical populations of inference to which descriptions of genetic diversity likely apply (i.e., the extent over which inferences applied) might therefore be loosely defined as: (1) naturally reproducing subpopulations of resident and anadromous *O. mykiss* affected by hydropower management on the Skagit River (e.g., hatchery supplementation, isolation, and hybridization with *O. clarkii*); and (2) artificially reproducing hatchery populations that are commonly used to supplement *O. mykiss* in the Skagit River basin. Due to the relatively short, two-year period over which sampling occurred, inferences might only reflect a "snapshot" of genetic diversity and could be limited to the one or two generations sampled between 2008 to 2010.

Because the Bull Trout samples analyzed in this memo were collected by multiple researchers, the scope of inference is harder to define than for Rainbow Trout. The study designs of Bull Trout described in Smith (2010), Small et al. (2013), and Small et al. (2016) were observational in nature. The most comprehensive of the three studies was completed by Smith (2010) in which 16

microsatellite markers were analyzed 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.*" Five objectives were listed, which can broadly be categorized as descriptions of genetic diversity and assignments of individuals from potentially mixed fishery collections to natal subpopulations of origin. Five-hundred-ninety-five juvenile and adult Bull Trout were sampled from fourteen localities due to their proximity to City Light's hydroelectric facilities and to represent populations' "baseline" localities likely to contribute to the genetic diversity of the adult Bull Trout found in the mainstem Skagit River. Samples were collected using a combination of electrofishing, snorkeling, and angling. No data were available regarding the extent to which sampling was opportunistic, targeted, or random with respect to age, phenotype, and life history. During 2006-2008, 435 fluvial adult and sub-adult Bull Trout were collected from the mainstem Skagit River from the Gorge Powerhouse to the confluence of the Sauk River. These samples represent a potential mixture from several spawning populations located throughout the Skagit basin. Samples were collected primarily by angling. The scope of inference for Bull Trout considered by this memo might therefore be defined as naturally reproducing subpopulations of Bull Trout in the Project vicinity that were affected by management of the Skagit Hydroelectric Project (e.g., isolation and hybridization) from 2005 to 2015.

2.2.1 Limitations of *Post Hoc* Studies

Due to the *post hoc* nature of analyses presented in this memo, statistical methods were chosen based on their ability to accommodate assumptions common to population genetic studies of **microsatellite** genotypes in general (words in bold appear in Glossary Section 6.0 of this memo). Typical guidance is to collect enough genetic samples in each **subpopulation** to accurately characterize **allele frequencies** (Landguth et al. 2010). For **iteroparous** species like Rainbow Trout and Bull Trout that display **overlapping generations**, sampling should at least be representative of the cohorts comprising the generation(s) of interest (Allendorf and Phelps 1980). A common assumption is that sampling occurs within predefined and discreet subpopulations and that **genetic population structure** occurs along a reproductive continuum ranging from panmixia (random mating) to complete isolation (mating is restricted to within subpopulations) (Waples and Gaggiotti 2006). Genetic diversity is often assumed to be a function of **genetic drift** and **gene flow** because microsatellites are assumed to be **selectively neutral** and the rate at which new genetic diversity enters the population through mutation is negligible (Waples and Gaggiotti 2006). Assuming further that there is sufficient **statistical power**, and absence of genotyping and sampling errors, estimates of genetic diversity are expected to reflect the true genetic structure of the population. In nature, all these assumptions are never met.

2.2.2 Genetic Analysis of Rainbow Trout

Due to the limitations described above, collections of genotypes that appeared to have been sampled from the same localities were pooled and treated them as random samples from individual "subpopulations" of Rainbow Trout unless statistical evidence that suggested they should be separated was observed (i.e., temporally spaced collections from the same locations were pooled unless the null hypothesis of genetic homogeneity was rejected). The computer program POWSIM Version 4.1 (Ryman and Palm 2006) was used to estimate statistical power to detect deviation from genetic homogeneity. POWSIM is a simulation-based computer program that estimates statistical power of rejecting the null hypothesis (H_0) of genetic homogeneity for different combinations of sample sizes, number of loci, number of alleles, and allele frequencies for a

hypothetical degree of true differentiation (quantified as F_{ST}). POWSIM can only accommodate 30 collections of individuals, so the first 30 collections (Table 1) were used to estimate power to detect low (F_{ST}=0.001) and moderate (F_{ST}=0.01) genetic differentiation by assuming allele frequencies estimated in Pflug et al. (2013). The program FSTAT Version 2.9.3.1 (Goudet 1995) was used to estimate and test metrics of genetic diversity unless otherwise stated. Expected heterozygosity (H_S, i.e., gene diversity assuming Hardy-Weinberg equilibrium [HWE]) and allelic richness (A_R) were estimated to describe genetic diversity across loci and collections. One thousand randomizations of alleles at each locus were performed to test the assumption of HWE at within collections. Observed (H_O) and expected multilocus heterozygosity within subpopulations were compared using Wright's (1951) F_{IS} to measure the magnitude of departures from HWE. To assess the assumption of random association of alleles among loci, log-likelihood ratio tests using 1,000 permutations were implemented to test for pairwise linkage disequilibrium (LD) within all collections. The Weir and Cockerham (1984) version of F_{ST} was estimated to measure genetic differentiation between all pairs of collections. To summarize genetic diversity among the sampled individuals, a principal component analysis (PCA) of individual-based genetic distances based on allele frequencies was implemented using the R package {adegenet} (Jombart et al. 2010). Ordination in "allelic space" along the first three PC axes was visualized using ggplot in program R. The statistical power to observe relatives was determined using {CKMRSim} (Anderson 2019). All tests of significance at the $\alpha = 0.05$ level were assessed and applied Bonferroni corrections when conducting multiple tests.

Table 1. Summary statistics for samples collected from *O. mykiss* in the Skagit and Fraser River basins.

Collection number [1]	Collection size	WDFW Code [2]	Location	Origin[3]	Upper Skagit[4]	Stage	Phenotype[5]	F_{IS} [6]	H_S [7]	A_R [8]	R^2 [9]
1	57	07MS, 08MI, 10BA	Bacon Creek	NOR	No	Juvenile, adult		0.01	0.79	9.45	0.02
X*	57	09EL	Baker River 09	NOR	No		Trout	0.09	0.82	10.44	0.03
X*	42	10AU	Baker River 10	NOR	No		Trout	0.11	0.84	11.13	0.04
2	51	09EU	Big Creek 09	NOR	No		Trout	0.04	0.66	5.63	0.02
3	48	10BG	Big Creek 10	NOR	No		Trout	0.06	0.67	5.16	0.03
4	52	09JB, 10BJ	Blackwater River	NOR	No	Juvenile	Trout	0.11	0.74	7.61	0.02
5	66	10MZ	Chilliwack Hatchery	HOR	No	Adult		0.00	0.76	8.02	0.02
6	94	09ET, 10BE	Clear Creek	NOR	No		Trout	0.06	0.68	8.56	0.01
7	38	10BB	County Line Ponds	NOR	No	Juvenile		0.01	0.80	9.10	0.04
8	26	05NG	Diablo	NOR	Yes		Trout	0.06	0.75	8.50	0.04
9	41	10BK	Diobsud	NOR	No	Juvenile		0.02	0.79	9.77	0.03
10	43	03OA	Dry Creek	NOR	Yes		Trout	0.02	0.71	7.50	0.03
11	47	09EH	Finney Creek	NOR	No	Juvenile		0.00	0.78	9.27	0.03
12	47	10AT	Finney Creek	NOR	No	Juvenile		0.01	0.80	9.59	0.02

Collection number [1]	Collection size	WDFW Code [2]	Location	Origin[3]	Upper Skagit[4]	Stage	Phenotype[5]	F_{IS} [6]	H_S [7]	A_R [8]	R^2 [9]
13	30	11BK	Finney Creek	NOR	No	Adult		-0.02	0.80	10.40	0.04
14	38	09IZ	Goodell Creek	NOR	No	Juvenile		0.01	0.77	8.76	0.03
15	41	10BC	Goodell Creek	NOR	No	Juvenile		0.00	0.79	9.12	0.03
16	47	09EE	lower Cascade	NOR	No	Juvenile		-0.05	0.77	8.23	0.03
17	44	10AV	lower Cascade	NOR	No	Juvenile		0.03	0.79	9.26	0.03
18	48	10AY	lower Skagit	NOR	No	Juvenile		0.02	0.79	9.51	0.03
19	28	08LF	lower Skagit	NOR	No	Adult		0.01	0.78	9.26	0.04
20	59	09CF	Marblemount	HOR	No	Adult		0.01	0.82	9.68	0.02
21	44	10AN	Marblemount	HOR	No	Adult		0.03	0.79	8.89	0.03
22	39	09BM	mid Skagit	NOR	No	Adult		0.01	0.80	10.49	0.03
23	31	10AS	mid Skagit	NOR	No	Adult		0.04	0.80	10.14	0.03
24	47	09ES	NF Cascade	NOR	No		Trout	0.11	0.41	4.30	0.02
25	45	10BF	NF Cascade	NOR	No		Trout	-0.08	0.36	3.98	0.02
26	79	02FB	Roland Creek	NOR	Yes		Trout	0.01	0.71	7.68	0.01
27	30	06AF	Ross	NOR	Yes		Trout	0.03	0.73	8.20	0.04
28	44	09MA	Ross	NOR	Yes		Trout	-0.01	0.69	6.65	0.04
29	47	10BH	Ross	NOR	Yes		Trout	-0.03	0.70	6.40	0.04
30	45	10AX	Sauk	NOR	No	Juvenile		0.04	0.80	9.66	0.03
31	29	83AAA	Sauk	NOR	No	Adult		0.06	0.80	10.29	0.04
32	32	09JA	Stetattle	NOR	Yes		Trout	0.03	0.76	8.66	0.04
33	41	10BI	Stetattle	NOR	Yes		Trout	0.03	0.77	8.79	0.03
34	115	09DT, 09EF, 10AQ, 10AW, 11BM	Suiattle	NOR	No	Juvenile, adult		0.01	0.79	10.05	0.01
35	51	09EV	upper Finney	NOR	No		Trout	0.03	0.74	6.52	0.02
36	49	10BD	upper Finney	NOR	No		Trout	0.04	0.72	6.77	0.02
37	56	10AZ	upper Skagit	NOR	No	Juvenile		0.01	0.79	9.56	0.02
38	32	11BI	upper Skagit	NOR	No	Adult		0.00	0.81	10.43	0.03

* Collections removed from PCA due to indirect evidence of hybridization with *O. clarkii*.

1 Collection number: corresponds to Figures 5 through 7.

2 WDFW code: WDFW collection identification with apparent sample year as the prefix

3 Origin: hatchery (HOR) or natural (NOR) origin.

4 Upper Skagit: collections from upstream of the Project Boundary in the Skagit River and from B.C.

5 Phenotype: identifies whether collections were from apparent trout as determined by WDFW.

6 F_{IS}: estimated deviation from HWE.

7 H_S: estimated expected heterozygosity within sub-populations (i.e., gene diversity).

8 A_R: estimated allelic richness.

9 R^2: is the estimated pairwise correlation of alleles among loci.

2.2.3 Genetic Analysis of Bull Trout

Exploratory analyses were conducted on Bull Trout similar to those described for Rainbow Trout. Partitioning of genetic variation was explored using visualization of individual-based data and genetic principal component analysis (e.g., Jombart et al. 2010). The statistical power to observe relatives was determined using {CKMRSim} (Anderson 2019). H_S was estimated following the sampling bias correction method described be Nei (1987). A common implementation of HWE test was used following Guo and Thompson (1992) Markov-chain random walk extension of Fisher's (2-allele) classical exact test. Departures from HWE were also quantified using the inbreeding coefficient (F_{IS}) statistic observed from analysis of molecular variance (AMOVA) (Excoffier et al. 1992; Yang 1998), which is equivalent to Weir and Cockerham (1984) small *f* statistics. Collections were analyzed for evidence of LD (i.e., non-independence of alleles at different loci). Given gametic phase was unknown for previously reported data, LD between a pair of loci was tested using a likelihood-ratio test, whose empirical distribution is obtained by a permutation procedure (e.g., Excoffier and Slatkin 1998). Lastly, allelic distributions across collections were evaluated using contingency table analysis of observed allelic distributions described by Raymond and Rousset (1995).

The AMOVA framework used to describe genetic structure of Bull Trout estimates hierarchical f-statistics for any number of desired levels (e.g., within individuals, within subpopulations, among subpopulations). This allows for subpopulation differentiation (allele frequency variance) to be quantified. In other words, the degree that individuals within a subpopulation (collection) are more similar to each other than are individuals from different subpopulations (collection). There are many formulations of the population differentiation variance component measure, although a common implementation is a form of the fixation index (e.g., genetic divergence [F_{ST}]). F_{ST} metrics were estimated pairwise following Weir and Goudet (2017) and used as a measure genetic divergence, with statistical significance calculated following likelihood-ratio tests (Goudet et al. 1996).

3.0 RESULTS

3.1 What genetic data are available?

3.1.1 Rainbow Trout

On 28 July 2021, Todd Seamons (WDFW) provided an Excel spreadsheet containing 2,967 genotypes for the 15 microsatellites analyzed in Pflug et al. (2013): *One-102*, *Ogo-4* (Olsen et al. 1998), *Ots-100* (Nelson et al. 1998), *Oki-10*, *Oki-23*(Smith et al. 1998), *Omy-7* (K. Gharbi, unpublished, as referenced in Pflug et al. 2013), *Omy-1001*, *Omy-1011*(Spies et al. 2005), *Ots-3M*, *Ots-4* (Banks et al. 1999), *One-14* (Scribner et al. 1996), *Ssa-407*, *Ssa-408* (Cairney et al. 2000), *Ssa-298* (McConnell et al. 1995), and *Oke-4* (Buchholz et al. 2001). The dataset included the following metadata: 'Sample Name', 'WDFW code', 'Count', 'Percent Missing Data'. Various other metadata were available directly from the Pflug et al. (2013) report (Table 1). The microsatellites appeared to be a subset of the standardized Stevan Phelps Allele Nomenclature (SPAN) markers described in Stephenson et al. (2009) that were developed to ensure data quality (repeatable allele scoring) across laboratories. Exact sampling locations (i.e., GPS coordinates) but based on the Pflug et al. (2013) report, appeared to include tributaries, mainstem rivers, hatcheries, and Project reservoirs. Some sites appeared to have been sampled across multiple years. The collections ranged in size from n=1 in the Suiattle River in 2009 to n=106 in Diablo Lake in 2005. No metadata were provided regarding sampling field methods (e.g., electrofishing), whether samples were collected randomly, or targeted life stages, life histories, morphologies, taxa, etc.

3.1.2 Bull Trout

On 13 June 2021, Matt Smith (USFWS) provided a tab delimited .txt file containing 563 genotypes at 16 microsatellite loci previously analyzed in Smith et al. (2010): *Omm1128*, *Omm1130* (Rexroad et al. 2001), *Sco102*, *Sco105*, *Sco106*, *Sco107*, *Sco109* (WDFW unpublished), *Sco200*, *Sco202*, *Sco212*, *Sco215*, *Sco216*, *Sco218*, *Sco220* (Dehaan and Ardren 2005), *Sfo18* (Angers and Bernachez 1996), and *Smm22* (Crane et al. 2004). The dataset included the following metadata: 'Individual Name', 'Synonym 1', 'Region (1)', 'Watershed (2)', 'Tributary (3)', 'Capture Location (4)', 'Age', 'Brood Year', 'Collected By', 'Collection Year', 'Comment', 'Date Collected', 'Fork Length (mm)', 'Hatchery/Wild', 'HOR/NOR assignment', 'Latitude', 'Life History', 'Stage', 'Longitude', 'Phenotypic Sex', 'PIT Tag', 'Population ID', 'Preservation Method', 'Project number', 'Received From', 'Resident / Anadromous', 'Run Type', 'Spawn Date', 'Spawn Year', 'Spawned With', 'Species', 'Synonym 2', 'Synonym 3', 'Tissue Type', 'Total Length (mm)', 'Used for Broodstock', 'Weight (g)'. Only some of these metadata were relevant to this report or contained entries. The same email sent from WDFW on 28 July 2021 contained 335 genotypes from six collections from the Project lakes and two collections from outside the Project Boundary.

3.2 What types of inference can be drawn from preexisting data?

3.2.1 Genetic analysis of Rainbow Trout

Of the 2,697 Rainbow Trout genotypes provided by WDFW, 536 were removed prior to analysis due to missing genotypes at two or more loci (e.g., as recommended by Reeves et al. 2016), and 20 were removed because their genotype was duplicated elsewhere in the dataset (i.e., they were removed due to possible pseudo replication). Pooling of samples from the same locations and across years reduced the number of analyzed collections from 76 to 25, however, only four of the

pooled collections contained fewer deviations from HWE and were thus retained: Bacon Creek (2007 to 2010), Clear Creek (2009 and 2010), Blackwater River (2009 and 2010), and the Suiattle River (1981 and 2009 to 2011). Collections with fewer than 25 individuals were removed to avoid biased estimates of allele frequencies within sub-populations as recommended by Hale et al. 2012. The provided table of genotypes did not contain any information about which markers were diagnostic for *O. clarkii* and so hybridization was not directly assessed. However, certain genotypes were removed from analyses if hybridization was indirectly apparent (i.e., if genotypes appeared to be statistical outliers).

Comparison of observed (H_O=0.729) and expected (H_S=0.747) heterozygosity across all collections and loci suggested a relatively small but overall deficit of heterozygotes (F_{IS}=0.025 95 percent CI [0.03, 0.01]). Eighty-six of 600 (14 percent) randomization tests for HWE (15 markers x 40 collections) were significant at the α=0.05 level with 68 (79 percent) of the tests showing a deficit of heterozygotes. No tests for HWE were significant at the adjusted level of α=0.00008. The locus *One-14* deviated from HWE in 17 of 40 (42.50 percent) total collections with all tests showing a deficit of heterozygotes. By contrast, most other markers (11 of 15) produced various combinations of heterozygote excess and deficiency. Therefore, the locus *One-14* was omitted from further analysis. This adjustment decreased mean F_{IS} to 0.017, though the difference was not statistically significant (95 percent CI: 0.03, 0.01).

The final dataset contained 1,900 individuals from 40 collections[2] genotyped at 14 microsatellites. However, in some instances, fewer than 40 collections and 1,900 individuals were analyzed (i.e., PCA) due to indirect evidence of hybridization with *O. clarkii*. The 14 microsatellite loci had a total of 312 alleles, ranging from 11 at *Ots-4* to 32 at *Omy-1001*. Across all 14 loci and 40 collections, the estimated false detection rate of a parent-offspring pair was 0.00000811, 0.00000033 of full siblings, and 7.277×10^{-21} of unrelated individuals. However, within any single collection, power is expected to be substantially lower. For example, the false positive rate (FPR) for related individuals in Roland Creek, a tributary within the Project Boundary, is 0.0000161 and the false negative rate is 0.392 (Figure 3). Gene diversity (H_S) within each collection ranged from 0.36 in the collection from North Fork Cascade River in 2010 to 0.83 in the Baker River in 2010 (Table 1). Average gene diversity in collections from upstream of the Project Boundary at Gorge Reservoir (H_S=0.74) was similar diversity in all other collections (H_S=0.74).

[2] While 40 collections were included in these study analyses, collections from the Baker River were removed from the PCA due to hybridization with *O. clarkii*.

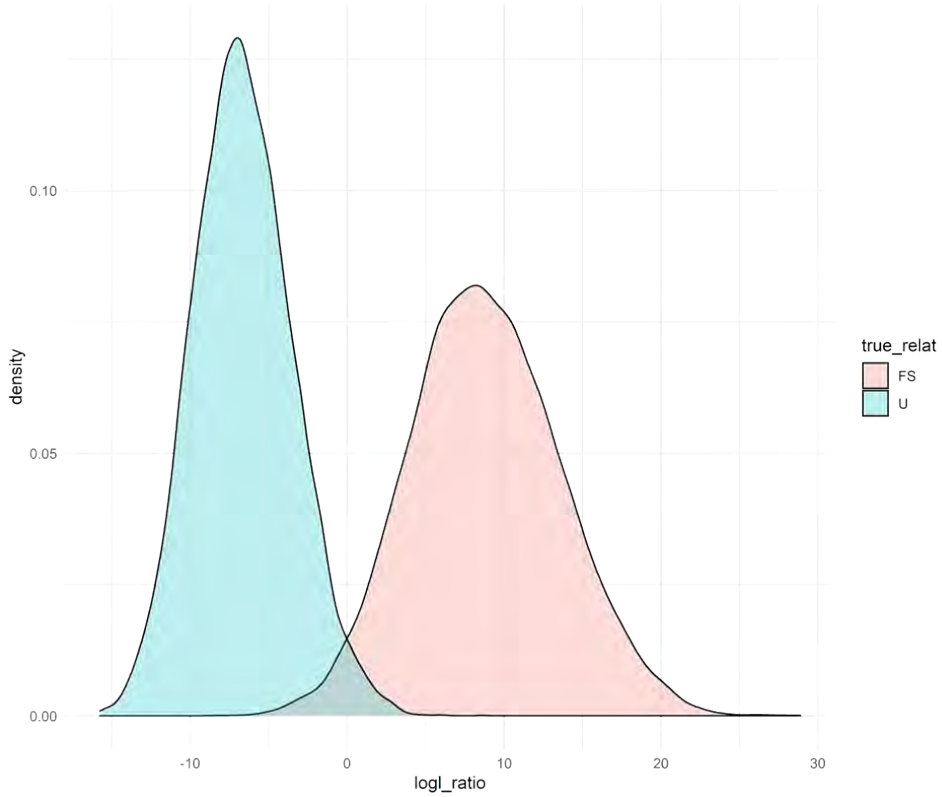


Figure 3. **Log-likelihood ratios distribution for simulated true full-siblings versus unrelated individuals based on Roland Creek *O. mykiss* genotype data. High overlap between full-siblings and unrelated fish, suggests relatively low power to detect highly related individuals.**

Six-hundred-forty of 3,640 (17.5 percent) log-likelihood (*G*) tests for pairwise LD using FSTAT were significant at the α=0.05 level. However, only 15 (<1 percent) tests were significant at the adjusted table-wide level of α=0.00007. The greatest disequilibrium was observed in the collection from Diablo Lake in 2005 (R^2=0.04) and the least in Suiattle River (R^2=0.01) (Table 1). Notably, there was a consistent, negative relationship between sample size and the estimator for pairwise LD, R^2 (Figure 4).

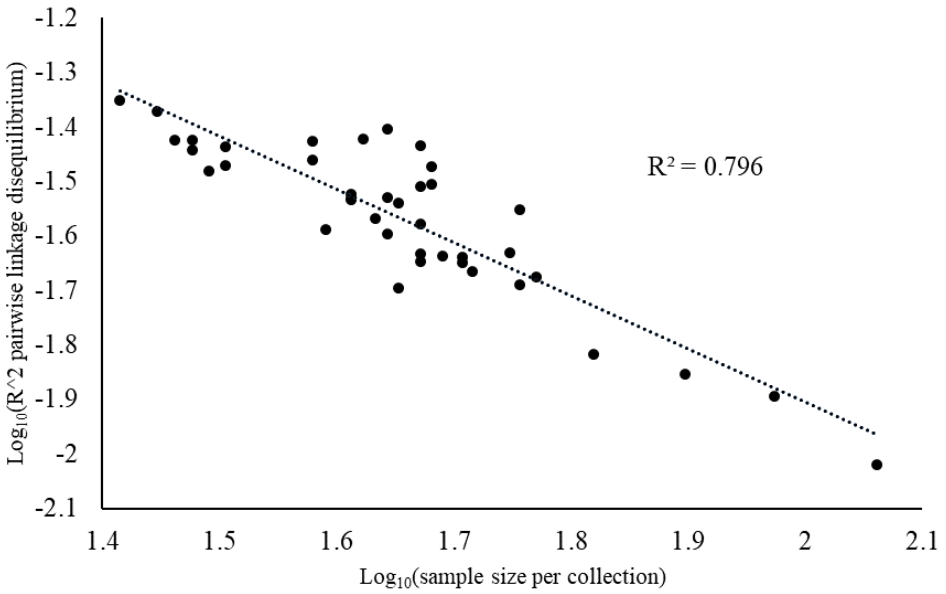


Figure 4. **Scatterplot showing $\log_{10}$-transformed relationship between sample size (n) (*x*-axis) and the R^2 estimator for pairwise linkage disequilibrium between loci (*y*-axis). Strong correlation warrants cautious interpretation of data due to possible bias.**

Fisher's exact tests using POWSIM (Ryman and Palm 2006) which were based on sample sizes and estimated allele frequencies of the dataset, suggested power to detect deviation from genetic homogeneity was 0.32 for F_{ST}=0.001 and was 1.00 for F_{ST}=0.01. The overall estimated proportion of genetic variance explained by population structure (F_{ST}) was 0.094. Log-likelihood (G) tests for population differentiation were significant for each locus and across all loci (P<0.001). Estimates of pairwise F_{ST} ranged from -0.004 between collections from Stetattle Creek in 2009 and 2010 to 0.39 between collections from Ross Reservoir and North Fork Cascade.

Principal components analysis (PCA) of allele frequencies using {adegenet} (Jombart 2011) accounted for a relatively small amount of projected inertia ─ a metric of the magnitude of the explained variance (Cumulative inertia explained by PC-1 through 3=5.924 percent). Genetic population structuring was apparent in scatterplots of the first three principal components (PC). However, several samples from the Baker River collections appeared to be outliers along axes 1 and 2. Notes provided by WDFW suggested the samples could be hybrids with *O. clarkii*. Reanalysis without the Baker River collections only slightly improved projected inertia of the first three PCs (6.095 percent); however, it did improve visualization of genetic population structure (Figures 5 and 6). Specifically, PC-1 (2.215 percent) clearly distinguished the North Fork Cascade River (Collections 24 and 25) from all other collections. PC-2 (2.044 percent) highlighted additional population structuring with collections from upstream of the project boundary at Gorge Reservoir tending to display positive inertia, collections from the Sauk River basin tending to display negative inertia, and remaining collections falling in between. PC-3 (1.836 percent) nearly distinguished Big Creek (Collections 2 and 3) from all other collections.

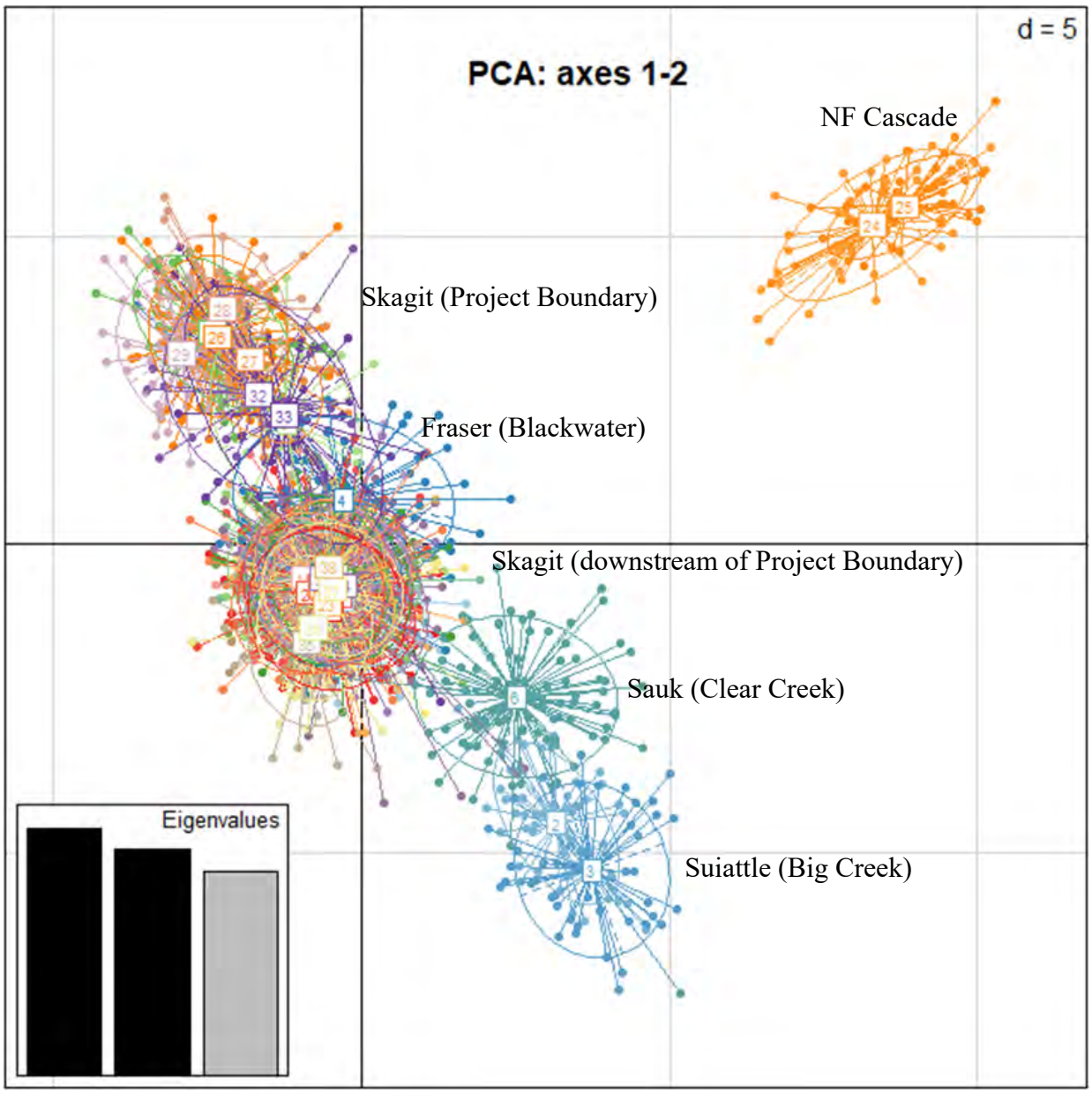


Figure 5. **Scatterplot of genetic principal components 1 (2.215 percent) and 2 (2.044 percent) for all Rainbow Trout collections, excluding samples from the Baker River, estimated using adegenet in program R (Jombart 2011). River basin names are provided to describe the approximate geographical locations of each collection. Numbers at centroids identify the collection number listed in Table 1. Ellipses define 1.5 standard deviations of the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom left corner shows first three eigenvalues.**

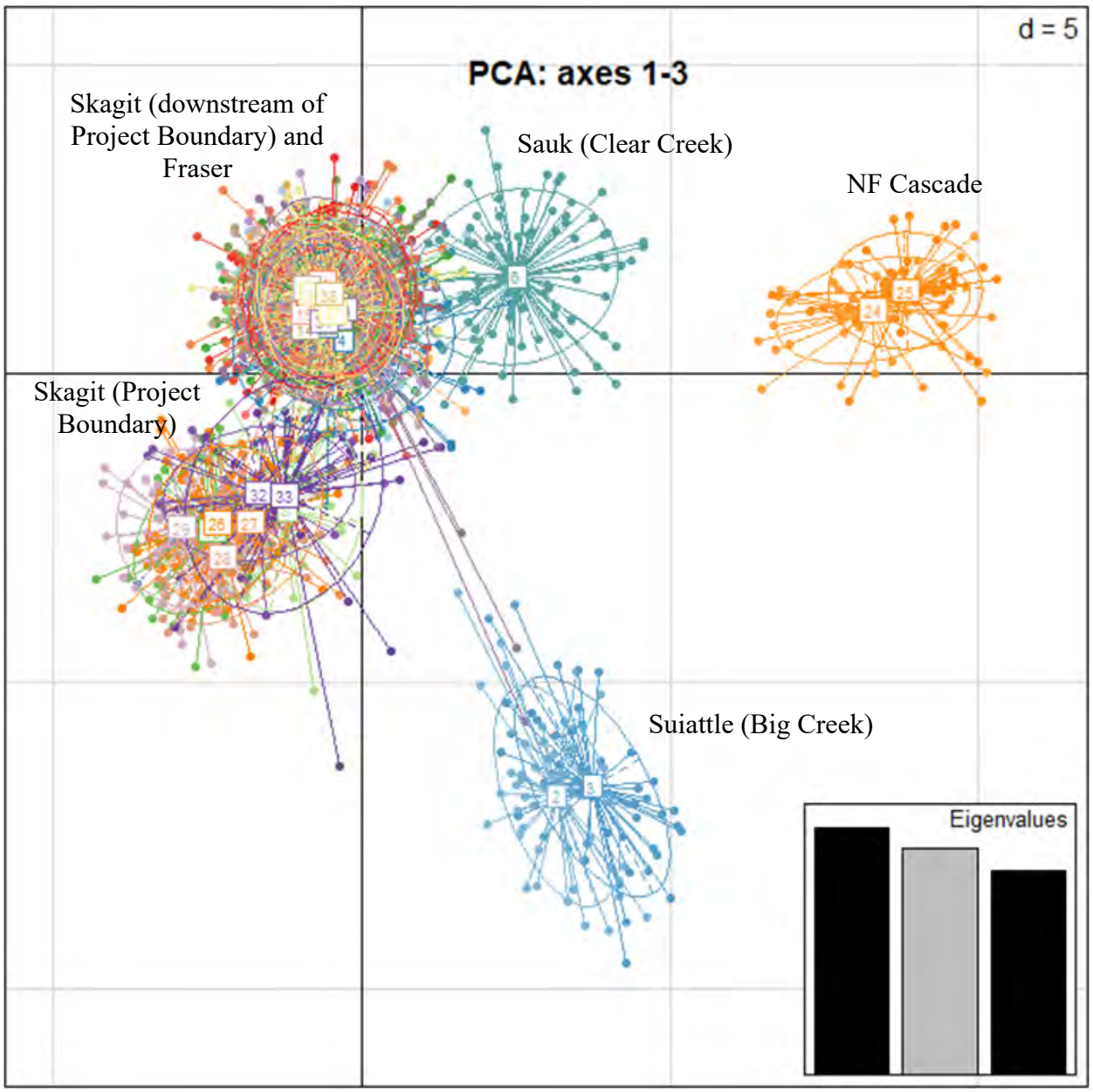


Figure 6. **Scatterplot of genetic principal components 1 (2.215 percent) and 3 (1.836 percent) for all Rainbow Trout collections, excluding samples from the Baker River, estimated using adegenet in program R. River basin names are provided to describe the approximate geographical locations of each collection. Numbers at centroids identify the collection number listed in Table 1. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom right corner shows first three eigenvalues.**

Limiting PCA to collections from upstream of the Project Boundary at Gorge Reservoir identified three samples that might be hybrids between *O. mykiss* and *O. clarkii* based on notes from WDFW; they were subsequently removed from the analysis (09JA0030, 05NG0056, and 10BI0047). Reanalysis without the potential hybrid samples indicated that the first three PCs explained 5.898

percent of the total inertia (Figure 7) and appeared to support some genetic structuring associated with location but statistical support for individual genetic groups was low.

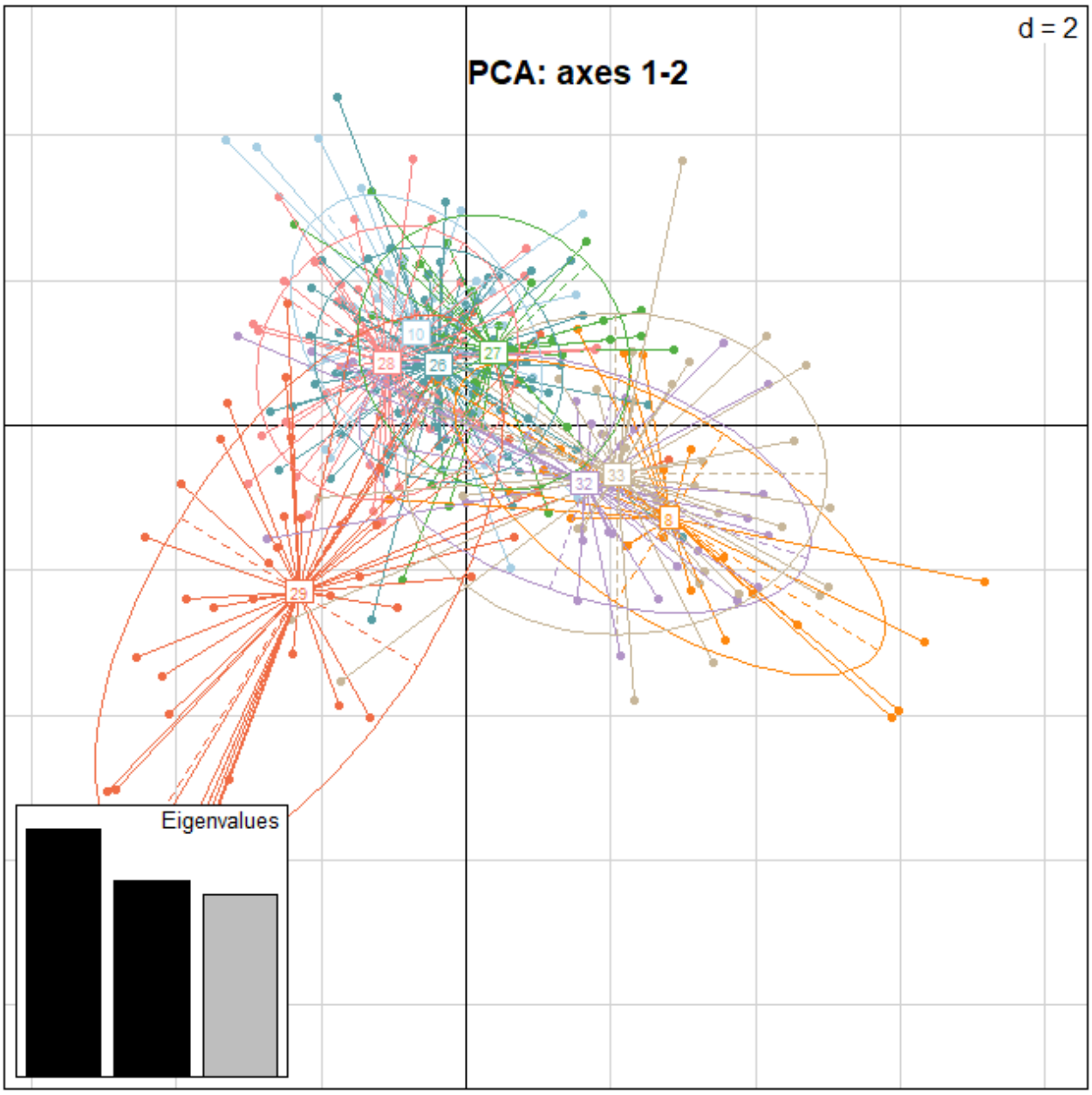


Figure 7. **Scatterplot of genetic principal components 1 (3.870 percent) and 2 (2.028 percent) for all Rainbow Trout collections upstream of the Gorge Lake Project Boundary estimated using adegenet in program R. Numbers at centroids identify the collection number listed in Table 1. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in bottom left corner shows first three eigenvalues.**

3.2.2 Genetic analysis of Bull Trout

Eight hundred and ninety-eight *Salvelinus* spp. genotypes at 16 microsatellite loci were provided by USFWS and WDFW following a request for existing Bull Trout data within Skagit River basin (Table 2). USFWS provided 563 of the genotypes and WDFW provided 335. The standardized markers included *Omm1128*, *Omm1130* (Rexroad et al. 2001), *Sco102*, *Sco105*, *Sco106*, *Sco107*, *Sco109* (WDFW unpublished), *Sco200*, *Sco202*, *Sco212*, *Sco215*, *Sco216*, *Sco218*, *Sco220* (Dehaan and Ardren 2005), *Sfo18* (Angers and Bernachez 1996), and *Smm22* (Crane et al. 2004). The collections were from four Project vicinity tributaries (upper Skagit, Big Beaver, Ruby, and Stetattle creeks) and all three reservoirs (Ross, Diablo, and Gorge lakes). It was unclear which *Salvelinus* taxa or their hybrids were included in the dataset. It was also unclear to what extent collections comprised highly related individuals, which is a common concern in genetic studies of Bull Trout (DeHaan et al. 2014). Furthermore, USFWS communicated that the juvenile collections likely contained related individuals (Smith 2021a).

Sampling location metadata were not provided for USFWS samples, so sampling locations were assumed to be the same as reported in Smith (2010). The stated purpose of the collections from Smith (2010) was to assess genetic variability within and between Bull Trout populations, with sampling methods including a combination of electrofishing, snorkeling, and angling.

No metadata were provided by WDFW other than collection code. Location data were not provided, so samples obtained from within the Project boundary were considered "at large" from reservoirs. The stated purpose of WDFW collections was to characterize the genetic variation of Bull Trout, Dolly Varden, and Brook Trout in the Skagit reservoirs, but no collection methodology was described. The degree to which samples were collected randomly across *Salvelinus* taxa was unknown, including whether any special effort was made to target Bull Trout, Dolly Varden, Brook Trout or whether potential hybrids were targeted or avoided. Sampling considerations are a key concern because targeted collections (i.e., based on morphology) can bias inference into studies of genetic variation.

Quality assurance/quality control (QA/QC) procedures were conducted to obtain a final dataset in which basic population genetic analyses could be reasonably implemented. Duplicate genotypes were observed for sample IDs 12FG008 and 12FG0009, and so sample 12FG0009 was omitted from dataset. All individuals with missing genotypes at three or more loci, which is more than the 14 loci chosen for *O. mykiss* (see above), were removed. This was necessary, however, because the Bull Trout data production appeared to have been conducted in four by four-locus panels (i.e., multiplexes), with many samples missing a single four locus block. Following data QA/QC, 589 samples were retained for analysis (Table 2).

Table 2. Bull Trout microsatellite dataset collection summary.

Collection Location	River	Life Stage	WDFW Code	Number Collected	Number Evaluated	Number Analyzed
Upper Skagit River	Skagit	adult		16	14	14
Big Beaver Creek	Skagit	adult		21	21	21
Ruby Creek	Skagit	adult		43	41	41
Stetattle Creek	Skagit	juvenile		59	41	41
Lower Goodell Creek	Goodell	juvenile		60	46	46
Upper Goodell Creek	Goodell	juvenile		19	8	8
Bacon Creek	Bacon	juvenile		61	24	24
Cascade River	Cascade	juvenile		39	33	33
Marble Creek	Cascade	juvenile		28	18	18
Kindy Creek	Cascade	juvenile		30	17	17
Illabot Creek	Illabot	juvenile		70	60	60
South Fork Sauk River	Sauk	juvenile		59	54	54
Downey Creek	Sauk	juvenile		58	44	44
Ross Lake	Skagit	unk	12FG	54	47	42
Ross Lake	Skagit	unk	15OW	28	22	20
Diablo Lake	Skagit	unk	13PS	40	29	8
Gorge Lake	Skagit	unk	14ST	27	5	3
Gorge Lake	Skagit	unk	19NL	109	22	0
Sulfur	Skagit	unk	050F	4	4	4
Sulfur	Skagit	unk	06JQ	28	23	23
Diablo, Gorge Lake	Skagit	unk	11LX	45	16	9
			Total	**898**	**589**	**530**

Statistical power was estimated to correctly classify related individuals. This was completed to evaluate the possible effects of violations of sampling assumptions common to the analysis of Bull Trout microsatellite data; specifically, that highly related individuals (i.e., full siblings) are common in samples of Bull Trout (particularly samples of juveniles), which can result in pseudo-replication of genotypes and thus biased estimates of allele frequencies (DeHaan et al. 2014). Statistical power of pedigree analysis to identify parent-offspring and full-sibling pairs was conducted using the close kin mark recapture R package CKMRSim version 0.1 (Anderson 2019; Formerly NMFS Southwest Fisheries Science Center). During pedigree analysis, all samples are examined for relatedness in pairwise comparisons, and so the false positive rate (FPR) increases exponentially with sample size. It is recommended to choose a FPR threshold approximately 10 times smaller than the reciprocal number of pairwise comparisons. In this case, $1.4\ e^{-5}$ was the target FPR used to evaluate the power to detect relatives (i.e., $0.10 * (100 \times 100)^{-1} = 0.000014$). To simulate the related and unrelated individuals needed to estimate power of pedigree analysis, all collections from the Skagit River dataset were used. The distribution of log-of-the-odds (LOD) values are shown in Figure 8 for full-sibling pairs. The expected distributions overlap between full-sibling and unrelated individuals, which means that choosing a FPR that provides reasonable assurance no unrelated pairs will be falsely called full-siblings will result in an undesirably high false negative rate (FNR). For Skagit River Bull Trout, a LOD value = 8.0 (corresponding to FPR

= 1.4 e^{-5}) results in a FNR = 0.15, meaning approximately 15 percent of true full-sibling comparisons would be misclassified as unrelated with an α=0.05 as the typical standard.

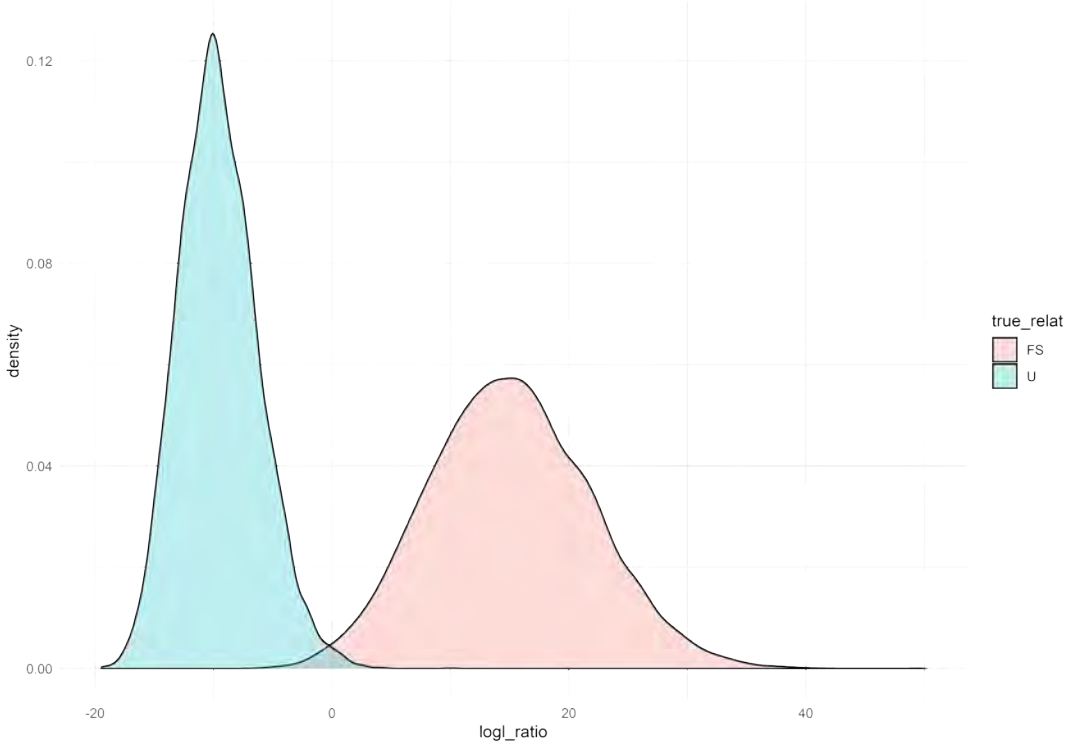


Figure 8. **Log likelihood ratios distribution for simulated true full siblings versus unrelated individuals based on Skagit River S. confluentus genotype data. High overlap between full-siblings and unrelated fish, suggests relatively low power to detect highly related individuals.**

Note that these estimated rates were based on all the available collections (n=530), which would likely overestimate power for studies of "real-world" populations. A more realistic evaluation would consider collections from a single Project Boundary tributary, as opposed to considering potential comparisons between unrelated individuals across the entire Skagit River basin. Therefore, the analysis was repeated, using only collections from Big Beaver, Ruby, and Stetattle creeks in the Project Boundary. The FNR estimated for Big Beaver, Ruby, and Stetattle collections were 0.857, 0.868, and 0.95, respectively, meaning pedigree analysis is expected to result in more false relationship assignments than true assignments.

Understanding power to detect related individuals helped identify individual samples that might need to be removed from analysis to reduce violation of sampling assumptions. COLONY (Jones and Wang 2010) was used to screen collections for full sibling families, and based on power estimates above, applied probability of inclusion = 1.0 and a probability of exclusion = 0.99 to accept family classifications. Inclusion probability gives the probability that all individuals (in that family) are indeed full siblings from the same family. Exclusion probability is the probability those individuals are full siblings, and no other individuals are full siblings with this family. There is no accepted convention or criterion for identifying and removing related individuals from a dataset, although the criteria used here are more stringent than those referenced in literature pertaining to

this Bull Trout dataset (Smith 2010). All full siblings but one[3] were omitted from identified families within collection.

Like for Rainbow Trout, PCA of allele frequencies (adegenet package) was used to examine genetic variation among collections. Data modeling suggested retention of approximately 15 PCs and 5 discriminant functions (*k*) would result in reliable partitioning of genetic variation among group clusters. With the number of genetic group clusters fixed at two (i.e., *k*=2), samples partitioned into genetic groupings associated with Diablo/Gorge lakes and all other samples. With an additional cluster allowed (*k*=3), individuals partitioned into (1) Project Boundary tributaries and some reservoir samples; (2) Project Boundary reservoir samples; and (3) samples from below Gorge Dam. With the allowance of fourth and fifth genetic clusters (*k*=4 and *k*=5), Project Boundary reservoir samples became split among the newly allowed clusters. No further refinement of Project Boundary samples was observed at higher numbers of clusters. A visualization of the k-means clustering at k=5 is shown on Figure 9. Clusters 3, 4 and 5 were predominantly individuals collected from Diablo and Gorge lakes. Cluster 1 were Project Boundary tributary collections and contained a majority of Ross Lake samples. Cluster 2 were individuals collected from below Gorge Dam.

As mentioned, collections submitted by WDFW were a part of evaluations intended to assess hybridization among Bull Trout, Dolly Varden and Brook Trout. Reports pertaining to data noted that hybrids were observed within these collections (e.g., Small et al. 2013; Small et al. 2016). Clusters 3, 4, and 5 were unable to be directly ascribed to hybridization among individuals or genetic introgression because: (1) taxon-diagnostic alleles among taxa were unknown; (2) sample IDs for individuals WDFW considered hybrids were not provided; (3) the methods by which WDFW determined individuals to be hybrids was not provided; and (4) the selection strategy (if any) of field personal collecting individuals "at large" from reservoirs was also not provided.

[3] The presence of multiple representatives from the same family skews allele frequencies from true population proportions, creating a bias. Removing all but one sibling removes this bias.

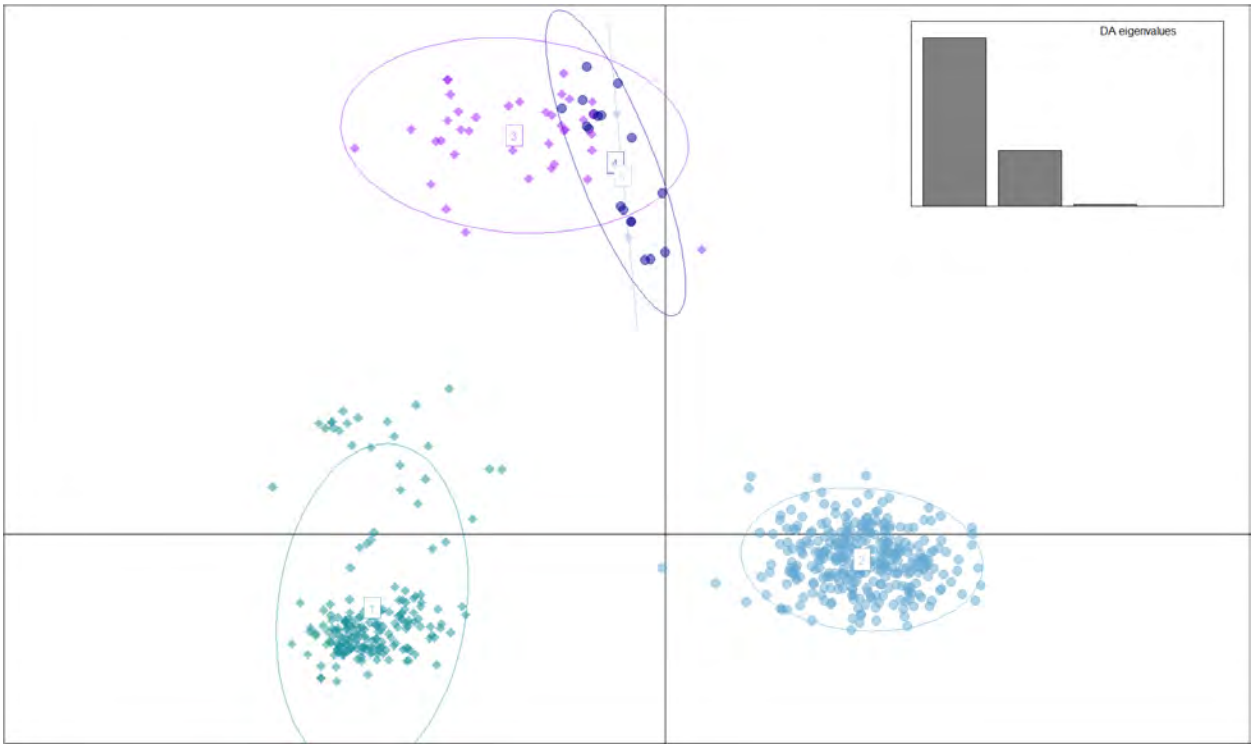


Figure 9. **Visualization of k-means clustering analysis at k=5 for Bull Trout individuals in dataset for 1st and 2nd principal axes. Ellipses define 1.5 standard deviations for the inertia (variance) around each centroid, where ellipses that overlap more are less distinct. Scree plot in upper right corner shows first three eigenvalues. Cluster 1 were Project Boundary tributary collections and contained a majority of Ross Lake samples. Cluster 2 were individuals collected from below Gorge Dam. Clusters 3, 4 and 5 were predominantly individuals collected from Diablo and Gorge Lakes.**

Small sample sizes of *Salvelinus* spp. (median=26) relative to *O. mykiss* (median=45) highlighted limitations associated with balancing precision and bias. For instance, collections with fewer than 25 individuals are typically not recommended for analyses using microsatellite data, however, adopting this criterion for the *Salvelinus* spp. dataset would have resulted in exclusion of about 50 percent of Project Boundary Bull Trout collections from an already sparse dataset. The genetic groupings shown in Figure 10 also underscore the challenges associated with choosing which fish to retain in any given collection due to genetic admixture. All individuals in clusters 3, 4, and 5 were considered potentially admixed and omitted from the dataset prior to estimating genetic summary statistics for each collection. The current sample size threshold pertaining to Bull Trout collections may be modified based upon future discussions of hypotheses and research questions with the Expert Panel and LPs. The resulting final dataset comprised n=530 samples (Table 2). The genotypes are saved in GENEPOP format and are available upon request.

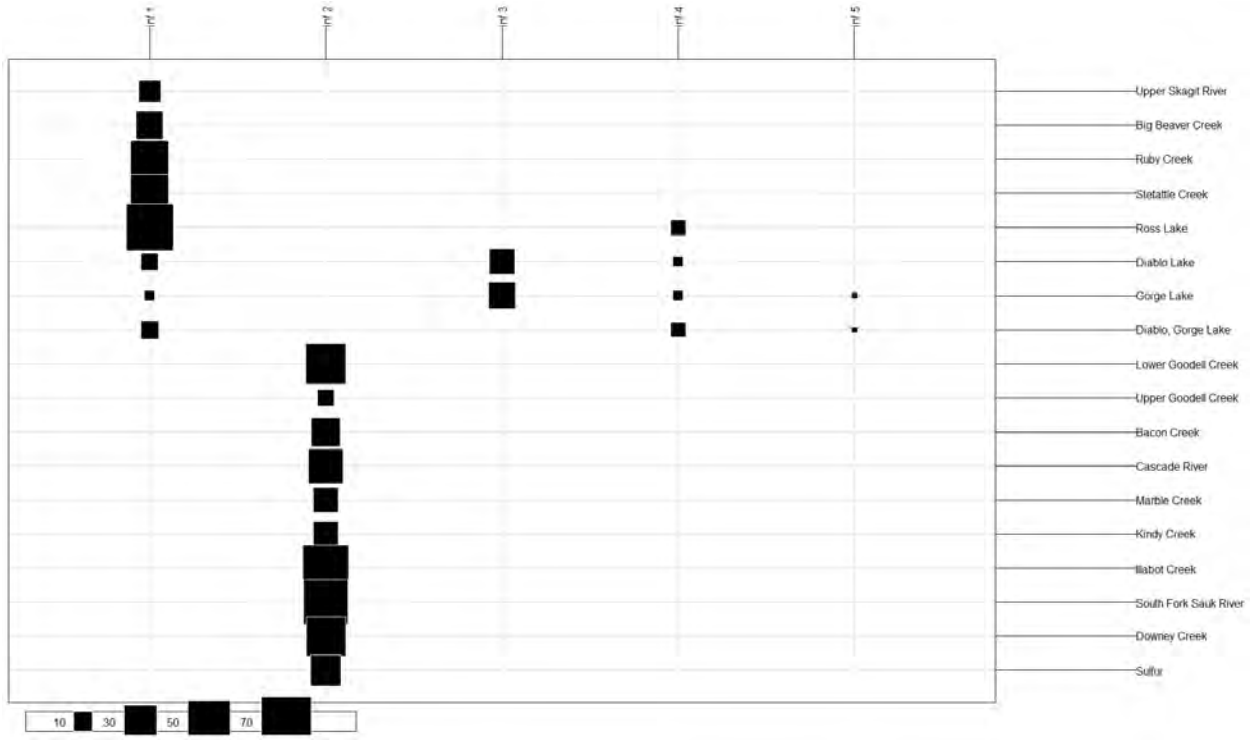


Figure 10. **Genetic clusters visualized in Figure 9 aligned to each Bull Trout collection in dataset. Size of boxes is scaled by sample count. Genetic clusters are organized by geographic location with upper Skagit collections at the top and lower Skagit at the bottom.**

Heterozygosity (gene diversity) in the Bull Trout collections ranged from 0.337 to 0.467 within collections from Project Boundary tributaries (above Gorge Dam) and was 0.473 in the Ross Lake collection (Table 3). The collections from within the Project Boundary (above Gorge Dam) had lower heterozygosity than the collections from below Gorge Dam (Chi-square p-value = 0.0027). Our attempt to reduce violation of HWE appeared successful, as mean F_{IS} across all collections was not statistically different from 0.00 (F_{IS}=0.008, 95 percent CI: -0.024-0.051). Each Project vicinity tributary collection (upper Skagit, Big Beaver, Ruby, Stetattle) did not deviate significantly from expectations. The Ross Lake collection was not in HWE, along with potentially several collections from below Gorge Dam, particularly Bacon Creek and Illabot Creek. LD was measured using log-likelihood (G) tests for all pairwise locus comparisons. Of the 1,680 comparisons (overall collections), 271 were significant at the α=0.05 level. No Project Boundary tributary collections (above Gorge Dam) had statistically significant LD tests using the adjusted table wide significance level α=0.0003. The Ross Lake collection had 11 significant LD test out of 120. The greatest number of significant log-likelihood tests was observed for the Illabot Creek collection (16).

The estimated proportion of genetic variance explained by population structure (F_{ST}) across all Bull Trout collections was 0.188, and 0.03 among Project Boundary tributary collections, only. Pairwise log-likelihood (G) tests for population differentiation were not statistically significant between upper Skagit, Big Beaver, and Ruby Creek collections (adjusted nominal level 5 percent). Upper Skagit River collection was not differentiated from Ross Lake collection, but Ross Lake collection was differentiated from both Big Beaver and Ruby Creek collections. The Stetattle

Creek collection was differentiated from all other Project Boundary collections. Note that the Marble Creek collection was not differentiated from any collection in the dataset except South Fork Sauk River. This seemed anomalous, so results that follow exclude consideration of Marble Creek collection. All Project Boundary collections (above Gorge Dam) were differentiated from below Gorge Dam collections. Recall, F_{ST} is the proportion of genetic variation that is attributable population subdivision with F_{ST}=0.00 reflecting no differences and F_{ST}=1.00 reflecting complete differentiation (i.e., all genetic diversity is partitioned among subpopulations). The F_{ST} estimated (pairwise) between Project Boundary collections are shown in Table 4). For context, F_{ST} estimated from comparisons between Project Boundary collections with those from below Gorge Dam ranged from a low of 0.207 to a high of 0.397.

Table 3. Summary statistics for samples collected from Bull Trout in the Skagit River basin.

Collection	Sample Size	F_{IS}[1]	H_s[2]	MNA[3]
Upper Skagit River	14	0.080	0.467	5.00
Big Beaver Creek	21	0.042	0.410	4.44
Ruby Creek	41	-0.021	0.384	4.75
Ross Lake	62	0.105	0.473	7.16
Stetattle Creek	41	-0.078	0.337	2.94
Goodell Creek	54	0.046	0.647	6.97
Bacon Creek	24	0.038	0.678	7.56
Illabot Creek	60	-0.050	0.634	7.44
Cascade River	33	0.033	0.662	8.19
Marble Creek	18	-0.080	0.679	6.94
Kindy Creek	17	0.016	0.689	7.19
S.F. Sauk River	54	-0.032	0.656	8.31
Downey Creek	44	0.010	0.709	9.88
Sulfur	27	0.035	0.607	6.13

1 $F_{IS:}$ estimated deviation from HWE proportions.
2 $H_{S:}$ estimated expected heterozygosity within sub-populations (i.e., gene diversity).
3 MNA: is the mean number of alleles observed over all loci.

Table 4. Table of pairwise estimates of F_{ST} from Project Boundary Bull Trout collections.

	Upper Skagit River	Big Beaver Creek	Ruby Creek	Ross Lake
Big Beaver Creek	0.001			
Ruby Creek	0.028	0.014		
Ross Lake	0.023	0.043	0.061	
Stetattle Creek	0.068	0.030	0.034	0.105

4.0 SUMMARY

4.1 Rainbow Trout

4.1.1 What genetic data are available?

Review of 2,697 preexisting microsatellite genotypes provided by WDFW highlighted opportunities and gaps to drawing new inference about population genetic characteristics of Rainbow Trout affected by FERC relicensing. Any new inference gleaned from these existing data is naturally limited by the design of the original research that estimated the genotypes. In general, inferences drawn from tens of presumably neutral microsatellites ─ in this case 14 ─ are naturally limited to basic descriptions of genetic diversity and population structure because of required analytical assumptions (see Section 2.1 of this memo). Likewise, new inference is also limited by the spatial, temporal, and ecological scope of the original sampling. The statistical populations to which any new inference applies should therefore be defined and agreed upon by the study team prior to determination of how this existing data addresses outstanding conservation questions or how new sampling might be most effective.

4.1.2 What types of inference can be drawn from the existing data?

The type of inference that can be drawn from analysis of the existing data could be limited to basic descriptions of genetic diversity and population structure. Genetic structure was apparent in the analyzed collections. The overall estimated proportion of genetic variance explained by population structure (F_{ST}) was 0.094, and the PCA appeared to provide some evidence that geography affects structure. Nevertheless, specific hypotheses about how current or historical geography affects structure were not tested (e.g., isolation-by-distance versus historical hydrogeological connectivity with the Fraser River). The proportion of variation that can be explained by hybridization with either *O. clarkii* or HOR fish were not directly addressed. Firstly, notwithstanding completely diagnostic makers, the set of microsatellites has limited power due to the number and diversity of markers (Vaha and Primmer 2006). Secondly, the question of hatchery introgression was addressed by Pflug et al (2013). Pflug et al (2013) used a liberal hybrid cut off threshold of 20 percent introgression from HOR fish and stated, "the juvenile collections showed the presence of presumptive [HOR] hybrids in all collection areas". All these factors are important to consider because observed patterns of diversity may not reflect natural genetic drift and gene flow within and among natural-origin *O. mykiss*.

4.1.3 Analytical considerations

Rainbow Trout genotypes analyzed in this memo were compiled in a way that attempted to reduce biases common to microsatellite datasets (See Section 2.1 of this memo). The approach attempted to increase any biologically meaningful signal by reducing noise associated with (1) hybridization with *O. clarkii*; (2) small sample sizes; (3) missing and erroneous data; and (4) violation of HWE and LD.

Absence of hybridization with *O. clarkii* is a common assumption of *O. mykiss* genetic structure analysis because many classical analyses assume that genetic variation is a function of effective population size (N_e) and migration (m) within a single taxon (mutation is assumed negligible). For example, the equation $F_{ST} \approx 1/(4N_em + 1)$ used to describe the strength of gene flow on genetic divergence assumes the subpopulations contributing migrants comprise only *O. mykiss*. Although

the dataset analyzed contained genotypes at genetic markers apparently diagnostic for *O. clarkii*, their diagnostic properties were unknown because positive control genotypes for nonhybridized *O. clarkii* were not provided. In practice, this limits the ability to estimate evolutionary relationships among subpopulations, which are typically assumed to be a function of genetic drift and gene flow within *O. mykiss*, as opposed to ongoing genetic introgression of alleles from *O. clarkii*.

Small sample sizes can result in imprecise estimates of allele frequencies and thus weak biological inference. There is no accepted threshold or rule for sample sizes because sampling needs vary by hypotheses, research questions, and marker types (Landguth et al. 2010). For the *O. mykiss* dataset, the recommendation of Hale et al. (2012) was adopted – those 25 individuals are typically enough to accurately estimate allele frequencies using microsatellites. Nevertheless, others have cautioned that when allelic diversity per population is high, as is the case with microsatellites, sampling effort may need to surpass 80–100 individuals to have a high probability of detecting low frequency alleles (Ott 1992: Seeb et al. 2007). Yet, other studies have reported that for isolated populations (n=8,000 individuals), 20 individuals genotyped at 6 microsatellites could produce an accurate allele frequency distribution (Siniscalco et al. 1999). For the present dataset, excluding collections of n<25 provided high power (1.00, $P<0.05$) to detect moderate differentiation ($F_{ST}=0.01$), but low power (0.32, $P<0.05$) to detect low differentiation ($F_{ST}=0.001$).

Like questions of sample size, there is no accepted threshold or rule for treating missing and erroneous microsatellite data. Using computer simulations, Reeves et al. (2016) estimated that for every 1 percent of missing genotypic data, 2 to 4 percent fewer correct population assignments can be expected. They recommended limiting the percentage of missing data to approximately 2 percent, unless a greater amount can be justified. Therefore, all individuals with missing data at two or more loci (approximately 6 percent), which was the most missing data that could be accommodated in a dataset of 15 microsatellites without allowing only individuals with complete genotypes to be included, were removed. Regarding genotyping errors, 1 to 2 percent fewer correct population assignments are expected for every percentage increase in genotyping error (Reeves et al. 2016). Although there are a variety of computer programs available to estimate the frequency of genotyping errors in a dataset, most techniques are based on conformance of genotypes to HWE proportions.

HWE and absence of LD are common assumptions of population genetic analyses for a variety of reasons. Metrics of HWE, for example, can provide insight to mating systems of populations (i.e., inbreeding) or to data quality problems like genotyping issues, overrepresentation of families, etc. The data compilation method for *O. mykiss* attempted to reduce violations of HWE that might result from data quality problems with the goal being to increase chances that metrics reflect the actual underlying mating system. The compilation method of removing markers with consistent deviation from HWE (i.e., *One-14*) and combining collections from the same tributaries but in different years that produced fewer deviations from HWE resulted in a dataset with a lower overall F_{IS} than the original dataset, though the decrease was not statistically significant. For clarity, lower F_{IS} suggests the compilation method succeeded in reducing deviations from HWE.

Regarding LD, a consistent negative relationship between sample size (n) and the estimator R^2, which could indicate that, on average, sampling was not sufficient to obtain unbiased estimates of LD, was observed. The potential bias presents a challenge to data interpretation. For example, the

collection from Diablo Lake contained the highest LD (R^2=0.04) but also one of the smallest sample sizes in the entire dataset (n=26). This presents a data interpretation challenge because the collection from Diablo Lake also contained apparent hybridization with *O. clarkii*, as noted by WDFW, which is expected to cause an increase LD associated with genetic admixture between genetically dissimilar populations. It is therefore uncertain whether high LD in Diablo Lake is associated with something biologically meaningful, like hybridization, or is simply an artifact of bias associated with small sample size.

4.1 Bull Trout

4.1.1 What genetic data are available?

Review of 898 microsatellite genotypes provided by USFWS and WDFW highlighted potential data gaps and opportunities to drawing new inference about population genetic characteristics of Bull Trout affected by FERC relicensing. Like Rainbow Trout, there are limitations to any new inference gleaned from these existing data. Most of the samples evaluated appear to have been previously analyzed by Smith (2010), Small et al. (2013), and Small et al. (2016) and so the collective scope of inference is somewhat ambiguous and should be discussed by the study group prior to making decisions on what questions can be answered by the existing data and what new samples need to be collected. Yet, it was uncertain which samples were evaluated in common among all three studies because sample identification were not included in the original reporting. The purpose for tissue sampling varied by collection. Smith (2010) stated that study's collection purpose was to assess genetic variability within and between Bull Trout populations of the Skagit River Basin and subbasins in the vicinity of City Light's Skagit Hydroelectric Project. The stated purpose of WDFW collections was to characterize the genetic variation of Bull Trout, Dolly Varden, and Brook Trout in the Skagit reservoirs. Given the sampling objectives differed for collections and multiple *Salvelinus* taxa (or hybrids) may have been incorporated into collections, it was challenging to compile a Bull Trout dataset. As with the *O. mykiss*, this approach focused on compiling the dataset that reduced violations of basic assumptions common to the analysis of microsatellites and, in general, to support basic inferences about genetic population structure of Bull Trout in the Skagit River basin. Nevertheless, as a working definition, the scope of inference of the analyses in this memo might apply to the few generations of naturally reproducing subpopulations of Bull Trout living within and downstream of the Project reservoirs and that might have been affected by key factors such as hydropower management (e.g., isolation) and hybridization from 2005 to 2015.

4.1.2 What types of inference can be drawn from the existing data?

Similar to Rainbow Trout, the types of inference that can be made about Bull Trout are likely limited to the few generations and subpopulations sampled by the original studies. The estimated proportion of genetic variance explained by population structure (F_{ST}) across all Bull Trout collections was 0.19. Project Boundary tributary collections were more similar to each other than any were to collections from below Gorge Dam, with collections from Ross Lake tributaries not statistically different. The Stetattle Creek collection was genetically differentiated from Ross Lake tributary collections. The "at large" Ross Lake collection was genetically differentiated from all Project Boundary collections except upper Skagit River. Pairwise F_{ST} estimates comparing Project Boundary collections with those from below Gorge Dam were relatively large in magnitude, with the minimum estimate observed being 0.207.

4.1.3 Analytical considerations

Like descriptions above for *O. mykiss*, evaluation of *Salvelinus* spp. data had to contend with inclusion of potential hybridized individuals, small collection sample sizes, missing and erroneous data, and violation to genetic equilibria. Data quality recommendations as noted above for Rainbow Trout were also applied to Bull Trout. The sample size threshold was reduced to retain collections from within the Project Boundary. Additionally, the missing genotype data threshold was increased to 25 percent and *Salvelinus* spp. samples that appeared ambiguous were omitted from summary statistic estimations. Lastly, while there was limited power to identify related individuals within collections, full-sibling families that were inferred using established methods were reduced in size. These steps resulted in a dataset that largely conformed to genetic equilibrium expectations, which was an improvement in data quality.

If genotypes at each genetic marker location (locus) occur at a frequency expected by random associations of alleles (a function of the allele frequency), genotypes are said to be in HWE, or alleles within loci are uncorrelated (statistically independent). Many phenomena may cause deviations for HWE expectations (e.g., null alleles, inbreeding, population mixing), with the deviation quantifiable using an analysis of variance approach. F statistics partition the reduction (or excess) in heterozygotes relative to HWE. One component, F_{IS}, is the individual relative to the subpopulation (collection). Globally across all collections, the mean F_{IS} observed was low (F_{IS} =0.008) and the 95 percent confidence interval overlapped zero. Further, all Project Boundary tributary collections were statistically consistent with HWE. LD quantifies the correlation of alleles between loci. LD is a useful quantity to measure, as the pattern of LD in the genome is influenced by population history, the breeding system, the pattern of geographic subdivision, natural selection, gene conversion, and mutation (Slatkin 2008). No Project Boundary tributary collections had statistically significant LD tests. The Ross Lake collection was not in HWE and had 11 statistically significant LD tests.

From a genotype frequency perspective, population structure results in an inbreeding like effect (a reduction in heterozygotes expected relative to HWE) due to nonrandom mating among all individuals analyzed. As such, measuring the deviation from HWE expectations due to population structure acts as a measure of genetic distance between two populations (or collections in this case). The degree of deviation is quantified using another F statistic, F_{ST}, the component of genetic variance within subpopulation (collection) relative to total population (paired collections being tested). For example, as mentioned above, Project Boundary tributary collections were more similar to each other than any were to collections from below Gorge Dam.

5.0 CONCLUSION

Existing microsatellite genotypes and metadata were available for Rainbow Trout (Pflug et al. 2013) and Bull Trout (Smith 2010; Small et al. 2013, 2016 and 2020) sampled within the Project Boundary and in the Project vicinity at different times and locations over the last twenty years. Genetic structure was apparent in both taxa and geography appeared to be an important factor in how the variation was distributed. Hypotheses about what created the structure were not tested. How structure was affected by hatchery introgression, hybridization, historical hydrogeological connectivity, genetic drift, etc. may be included in Expert Panel discussions and FA-06 Reservoir Fish Genetics Study reporting. Further, whether these data are sufficient to inform topics of interest communicated by LPs will be considered as part of Expert Panel discussions when topics are transformed to specific scientific question that can be applied to these data. Additionally, there are other genetic datasets that have been collected and analyzed by researchers that were not considered in this memo. As stated above, the data analyzed in this tech memo were recognized as potentially helpful for describing the baseline genetics of native fish in the RSP. Yet, unconsidered data could be brought to bear during formulation of research questions or designs for additional field sampling. Datasets not included here include genotypes for different genetic markers, including mitochondrial DNA haplotypes for Bull Trout inside of the Project Boundary (Smith 2021b) and for outside of the Project Boundary (Taylor and May-McNally 2015). There are also single nucleotide polymorphism (SNP) genotypes for 30 Rainbow Trout sampled from the drawdown zone of Gorge Lake (Small et al. 2020).

6.0 GLOSSARY

Allele frequency: a measure of the relative frequency of an allele at a genetic locus in a population.

genetic drift: random changes in allele frequencies in a population between generations due to sampling individuals that become parents and binomial sampling of alleles during meiosis.

Gene flow: exchange of genetic information between subpopulations.

Genetic population structure: systematic difference in allele frequencies between subpopulations in a population resulting from non-random mating between individuals.

Iteroparous: a reproductive strategy characterized by multiple reproductive cycles over the course its lifetime.

Microsatellite: tandemly repeated DNA consisting of short sequences of 1 to 6 nucleotides repeated approximately 5 to 100 times.

Overlapping generations: a breeding system where sexual maturity does not occur at a specific age, or where individuals breed more than once, causing individuals of different ages to interbreed in a given year.

Selectively neutral: an allele that is not under selection because it has no effect on fitness.

Statistical power: probability of obtaining a statistically significant result given a true effect occurs in a population.

Subpopulations: groups of individuals within a population delineated by reduced gene flow with other groups.

7.0 REFERENCES

Allendorf, F. W., and S. R. Phelps. 1981. Use of allelic frequencies to describe population structure. Canadian Journal of Fisheries and Aquatic Sciences 38:1507-1514.

Anderson, S. 2019. CKMRsim package, efficient sampling for close-kin mark recapture power estimation. Available at https://github.com/eriqande/CKMRsim

Angers, B. and L. Bernachez. 1996. Usefulness of heterologous microsatellites obtained from brook charr, *Salvelinus fontinalis* Mitchill, in other *Salvelinus* species. Molecular Ecology 5:317–319.

Banks, M. A., M. S. Blouin, V. K. Baldwin, H. A. Rashbrook, H. A. Fitzgerald, S. M. Blankenship, and D. Hedgecock. 1999. Isolation and inheritance of novel microsatellites in Chinook Salmon (*Oncorhynchus tschawytscha*). The Journal of Heredity 2.

Buchholz, W. G., S. J. Miller, and W. J. Spearman. 2001. Isolation and characterization of chum salmon microsatellite loci and use across species. Animal Genetics 32.

Cairney, M., J. B. Taggart, and B. Hoyheim. 2000. Characterization of microsatellite and minisatellite loci in Atlantic salmon (*Salmo salar* L.) and cross-species amplification in other salmonids. Molecular Ecology 9:2155-2234.

Crane, P. A., C. J. Lewis, E. J. Kretschmer, S. J. Miller, W. J. Spearman, A. L. DeCicco, M. J. Lisac, and J. K. Wenburg. 2004. Characterization and inheritance of seven microsatellite loci from Dolly Varden, *Salvelinus malma*, and cross-species amplification in Arctic char, S. alpinus. Conservation Genetics 5:737-741

DeHaan P. W. and W. R. Ardren. 2005. Characterization of 20 highly variable tetranucleotide microsatellite loci for bull trout (*Salvelinus confluentus*) and cross amplification in other Salvelinus species. Mol. Ecol. Notes 5: 582-585.

DeHaan, P., B. J. Adams, and M. C. Nelson. 2014. Fine-Scale Population Structure Analysis and Genetic Population Assignments of Wenatchee River Sub-Basin Bull Trout.in D. o. t. I. U. S. F. W. Service, editor., Longview, WA.

Excoffier, L., and M. Slatkin. 1998. Incorporating genotypes of relatives into a test of linkage disequilibrium. Am. J. Hum. Genet. 171-180.

Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Goudet, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of Heredity 86:485-486.

Goudet, J., M. Raymond, T. de Meeüs, and F. Rousset. 1996. Testing differentiation in diploid populations. Genetics 144: 1933-1940.

Guo, S. and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48:361-372.

Hale, M. L., T. M. Burg, and T. E. Steeves. 2012. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population Is enough to accurately estimate allele

frequencies. PLoS ONE 7:1-10.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal 732 components: a new method for the analysis of genetically structured populations. BMC 733 Genet. 11:94. BioMed Central Ltd.

Jombart, T. 2011. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jones, O. R., and J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10:551-555.

Landguth, E. L., S. A. Cushman, M. K. Schwartz, K. S. McKelvey, M. Murphy, and G. Luikart. 2010. Quantifying the lag time to detect barriers in landscape genetics. Molecular Ecology 19:4179–4191.

McConnell, S., L. Hamilton, D. Morris, D. Cook, D. Paquet, P. Bentzen, and J. Wright. 1995. Isolation of salmonid microsatellite loci and their application to the population genetics of Canadian east coast stocks of Atlantic salmon. Aquaculture 137:19-30.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY, USA.

Nelson, R. J., T. D. Beacham, and M. P. Small. 1998. Microsatellite analysis of the population structure of a vancouver island sockeye salmon (*Oncorhynchus nerka*) stock complex using nondenaturing gel electrophoresis. Molecular Marine Biology and Biotechnology 7:312-319.

Olsen, J. B., P. Bentzen, and J. E. Seeb. 1998. Characterization of seven microsatellite loci derived from pink salmon. Molecular Ecology 7:1087-1089.

Ott, J. 1992. Strategies for characterizing highly polymorphic markers in human gene mapping. American Journal of Human Genetics 51:283-290.

Pflug, D. E., E. J. Connor, B. Hayman, T. W. Kassler, K. Warheit, B. McMillan, and E. Beamer. 2013. Ecological, genetic, and productivity consequences of interactions between hatchery- and natural-orgin steelhead of the Skagit watershed. Skagit River System Cooperative.

Queller, D.C. and K.F. Goodnight. 1989. Estimating Relatedness Using Genetic Markers. Evolution 43(2):258–275.

Raymond M. and F. Rousset. 1995. An exact test for population differentiation. Evolution 49:1280-1283

Reeves, P. A., C. L. Bowker, C. E. Fettig, L. R. Tembrock, and C. M. Richards. 2016. Effect of error and missing data on population structure inference using microsatellite data. BioRxiv. 080630. Available online at https://www.biorxiv.org/content/ early/2016/10/13/080630.

Rexroad, C. E. III, R. L. Coleman, A. M. Martin, W. K. Hershberger, and J. Killefer. 2001. Thirty-five polymorphic microsatellite markers for rainbow trout (*Oncorhynchus mykiss*). Animal Genetics 32:317-319.

Rousset, F. 2020. Genepop Manual. Version 4.7.5.

Ryman, N., and S. Palm. 2006. POWSIM: a computer program for assessing statistical power

when testing for genetic differentiation. Molecular Ecology Notes 6:600-602.

Scribner, K. T., J. R. Gust, and R. L. Fields. 1996. Isolation and characterization of novel microsatellite loci: cross species amplification and population genetic applications. Canadian Journal of Fisheries & Aquatic Sciences 53.

Seattle City Light (City Light). 2021. Revised Study Plan (RSP) for the Skagit River Hydroelectric Project, FERC Project No. 553. April 2021.

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

Siniscalo, M., R. Robledo, P. K. Bender, C. Carcassi, L. Contu, and J. C. Beck. 1999. Population genomics in Sardinia: a novel approach to hunt for genomic combinations underlying complex traits and diseases. Cytogenetic and Genome Research 86:148-152.

Slatkin, M. 2008. Linkage disequilibrium — understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 9: 477–485.

Small, M.P., E. Martinez, and S. Peterson. 2013. Genetic analysis of native char collected in Diablo Lake, Ross Lake, Gorge Reservoir, Skagit mainstem, and Skagit Bay. WDFW Molecular Genetics Lab, Olympia WA.

Small, M.P., S. Bell, and C. Bowman. 2016. Genetic analysis of native char collected in Diablo Lake, Gorge Reservoir, and Ross Lake in the Skagit River basin. Washington Department of Fish and Wildlife Molecular Genetics Lab, Olympia, Washington.

Small, M., E. Lowery, and M. Kissler. 2020. Genetic analysis of Char from Gorge Lake, Upper Skagit River, Washington. Final Report.

Smith, C. T., B. F. Koop, and R. J. Nelson. 1998. Isolation and characterization of coho salmon (*Oncorhynchus kisutch*) microsatellites and their use in other salmonids. Molecular Ecology 7:1614-1616.

Smith, M. J. 2010. Population structure and genetic assignment of bull trout (*Salvelinus confluentus*) in the Skagit River Basin. School of Aquatic and Fishery Sciences University of Washington.

_____. 2021a. Personal communication between Scott Blankenship, Science Director Cramer Fish Sciences and Matt Smith, Fish Geneticist USFWS. June 14, 2021.

_____. 2021b. Personal communication between Dan Bingham, Senior Geneticist Cramer Fish Sciences and Matt Smith, Fish Geneticist USFWS. December 28, 2021.

Spies, I. B., D. J. Brasier, and P. T. L. O'Reilly, T. R. Seamons, and P. Bentzen. 2005. Development and characterization of novel tetra-, tri-, and dinucleotide microsatellite markers in rainbow trout (*Oncorhynchus mykiss*). Molecular Ecology Notes 5:278-281.

Stephenson, J. J., M. R. Campbell, J. E. Hess, C. C. Kozfkay, A. Matala, M. V. McPhee, P. Moran, S. Narum, M. M. Paquin, O. Schlei, M. P. Small, D. M. Van Doornik, and J. Wenburg. 2009. A centralized model for creating shared, standardized, microsatellite data that

simplifies inter-laboratory collaboration. Conservation Genetics 10:1145-1149.

Taylor, E. and May-McNally, S. L. 2015. Genetic analysis of Dolly Varden (*Salvelinus malma*) across its North American range: evidence for a contact zone in southcentral Alaska. Canadian Journal of Fisheries and Aquatic Sciences. 72:1-10.

Vaha, J. and C. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63–72. doi:10.1111/j.1365-294X.2005.02773.x.

Waples, R., and O. E. Gaggiotti. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15:1419–1439.

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

Weir, B.S. and J. Goudet. 2017. A Unified Characterization of Population Structure and Relatedness. Genetics 206(4):2085–2103.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics. 15:323–354.

Yang, R.C. 1998. Estimating Hierarchical F-Statistics. Evolution 52(4):950–956.

RESERVOIR FISH GENETICS STUDY

ATTACHMENT B

YEAR 2 SAMPLING PLAN

Technical Memorandum

Date:	Friday, April 15, 2022
Project:	Skagit River Hydroelectric Project
To:	Seattle City Light
From:	Scott Blankenship and Dan Bingham, Cramer Fish Sciences
Subject:	FA-06 Reservoir Native Fish Genetics Baseline Study – 2022 Proposed Year 2 Sampling Plan

1.0 INTRODUCTION

Seattle City Light (City Light) owns and operates the Skagit River Hydroelectric Project (Project). City Light operates the Project under license issued by the Federal Energy Regulatory Commission (FERC) and is in the process of pursuing a new license (i.e., relicensing) with expiration of the current FERC license in 2025. To support the relicensing process, City Light is conducting a suite of environmental studies, including FA-06 Reservoir Native Fish Genetics Baseline Study (Reservoir Fish Genetics Study). City Light has been conducting the two-year Reservoir Fish Genetics Study in consultation with agencies, Tribes, and other interested parties (hereafter called licensing participants [LPs]).

2.0 GOALS AND OBJECTIVES

As described in the Revised Study Plan (RSP), the goal of the Reservoir Fish Genetics Study is to characterize baseline population genetic structure for three native salmonid species: Bull Trout, Rainbow Trout, and Dolly Varden (target species), in Project reservoirs and provide the basis necessary to inform the planning of long-term (i.e., over the next Project license term) reservoir fish management objectives. Specific goals include:

1. Determine the population genetic structure of within and among target species populations and assess whether management actions are necessary for genetic sustainability.
2. Determine the number of fish populations, for each target species, within and among the Project reservoirs.
3. Estimate the effective population size (N_e) for each target species and reservoir.
4. Identify topics and/or management objectives to be considered in the reservoir fish and aquatics management plan.

On June 9, 2021, City Light filed a "Notice of Certain Agreements on Study Plans for the Skagit Relicensing" (June 9, 2021 Notice) that detailed additional modifications to the RSP agreed to between City Light and supporting LPs. The June 9, 2021 Notice included the following additional Reservoir Fish Genetics Study objectives:

1. Collect juvenile fish at spawning grounds for genetics baseline as part of field sampling program in Year 2.

2. Expand sample collection and/or coordination of existing samples and activities and analysis out-of-basin and above/below dams.

Following discussions with LPs in the Fall 2021, City Light agreed to consider additional genetic management questions developed by LPs for potential integration into the Reservoir Fish Genetics Study. These questions have the potential to modify the scope of the Year 2 field program beyond the original RSP and June 9, 2021 Notice objectives and/or agreements identified above. Note that City Light is open to discussing but has not committed to adopting/integrating LP management questions as additional study objectives or as future Project-related management activities. Some management questions may not be feasible to implement as part of the Reservoir Fish Genetics Study (given the study schedule within the relicensing timeframe) and may be more appropriate as long-term research activities or part of future reservoir fisheries management objectives.

The questions submitted by LPs are as follows:

1. What is the genetic relationship of *O. mykiss* (OM), Bull Trout (BT), and Dolly Varden (DV) above and below each of the dams and how distinct are the fish confined by the dams in each watershed?
2. How are Upper Skagit populations of OM, BT, and DV genetically related to other regions/core areas in Washington State, British Columbia, and the PNW? Are there other or more appropriate spatial areas that should be considered?
3. How many genetically identifiable populations of OM, BT, and DV exist upstream of Gorge Dam? What are the genetically identifiable populations of OM, BT, and DV that exist below Gorge Dam? What genetic diversity (e.g., N_e) exists within these populations? Are the populations of BT, OM, and DV confined behind each dam viable and maintaining natural levels of genetic diversity or do population sinks exist?
4. How can we use genetic data to evaluate the amount of fish entrainment through spill and the penstocks? Is there enough data to identify how many populations are affected? Or can eDNA be used with water or even sediment cores to identify current populations affected or entrained?
5. Can we use genetic data to identify the migration timing (i.e., winter or summer run) of Skagit Basin O. mykiss? Are there non-genetic methods to determine migration timing, especially of resident populations? If yes, what would these methods be and how would a study be developed? Especially for "resident" trout populations migrating during drawdown periods.
6. To what degree, if any, is hatchery introgression affecting the OM in each reservoir? What is the genetic relationship of OM in each reservoir to the Ross Broodstock and how distinct is this relationship?
7. Can we use genetic data to identify species hybrids (i.e., BT, DV, & EBT and OM & CT)? At what resolution (e.g., F1, F2, backcross)? If so, can we determine the frequency and spatial extent of hybridization in the upper Skagit Basin? Can we determine the impact the dams and the ongoing operations, maintenance, and other management activities have had on hybridization? Can we identify hot spots for management of EBT?

Following receipt of LP research questions, the Expert Panel determined the need for refinement of the LP research questions. The timing for completion of this task has not yet been defined.

Completing a Year 2 field program in 2022 is critical to remaining on the study schedule. Study activities include finalizing the field sampling plan, acquire permitting approvals, completing the Year 2 sampling program, analyzing data, consultation with LPs and Expert Panel members, and reporting within the next 8 months. Given the ongoing discussions between the Expert Panel and LPs regarding the refinement of

LP questions and relevancy to the Project scope and City Light interests, this sampling plan was developed to achieve goals of the RSP and the June 9, 2021 NOA. Nevertheless, initial review of the LP questions (as described above) suggests this sampling plan will provide information useful (in part or entirely) for addressing several LP questions, once finalized.

In consideration of the above, the purpose of this memo is to describe the field, laboratory, and analytical methods to be implemented as part of the Year 2 field program to meet the objectives of the FA-6 Reservoir Fish Genetics Study (inclusive of the NOA commitments and where appropriate, in support of LP questions).

3.0 STUDY AREA

The study area associated with field sampling generally encompasses the Project reservoirs and associated tributaries in the U.S. (Figure 1). Although the geographic distribution of all three taxa extends across the international border into Canada, sampling across the border presents numerous logistical challenges due to permitting and the Covid-19 pandemic. Therefore, for the purposes of 2022 field sampling for the Reservoir Genetics Study, sampling will be conducted in the U.S. only. Additional coordination to meet out-of-Project area study objectives are not identified in the study area and will require additional consultation with LPs.

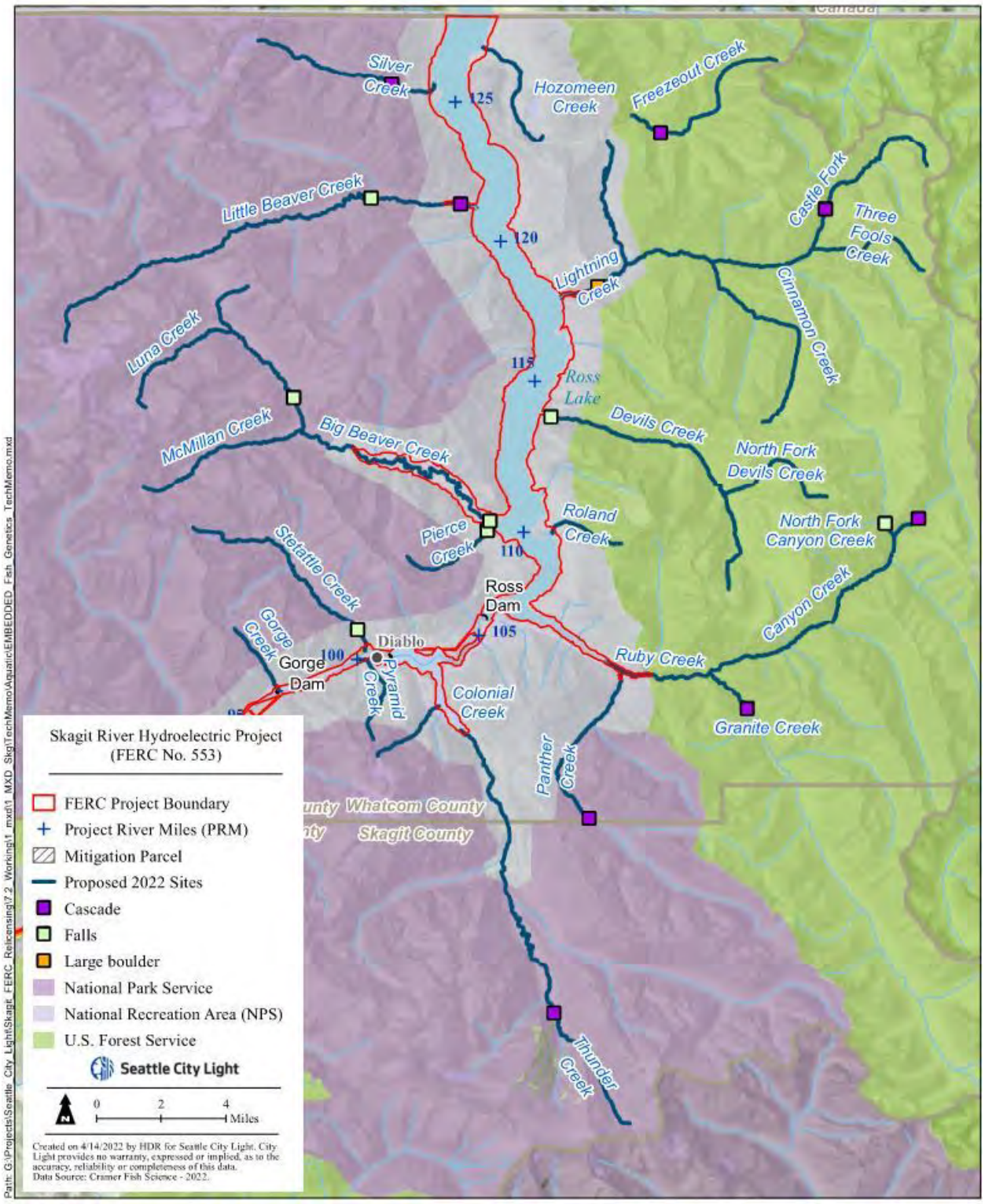


Figure 1: Map of Project Boundary and proposed sampling locations in the United States.

4.0 METHODS

4.1 Approach

In 2022, City Light will focus efforts on obtaining the samples and genotypes needed to characterize genetic population structure and relationships and to begin estimating N_e (to evaluate population viability). Accuracy and precision of the genetic parameters associated with these metrics are substantially affected by sample size and the number, diversity, and type of genetic marker (e.g., single nucleotide polymorphism [SNP] versus microsatellite). City Light will take advantage of newly developed GT-seq SNP markers, which include hundreds of SNPs useful for describing genetic diversity, hybridization, and some quantitative genetic traits, such as migration timing. The chosen SNPs are standardized and commonly genotyped by Washington Department of Fish and Wildlife (WDFW) and U.S. Fish and Wildlife Service (USFWS), which will provide a significant technological update to the current genetic database (i.e., microsatellite genotypes from Pflug et al. [2013]; Smith [2010]) and allow for straightforward comparison of data among laboratories. Sampling and genotyping in 2022 and the relevant study objectives are as follows:

1. Obtain tissue samples and genotypes for all target species to support:
    a. Determining the number of fish populations for each target species within and among Project reservoirs (RSP Objective 2)
    b. Describing population structure for Bull Trout and Rainbow Trout. Objective facilitated by using SNP genotypes (RSP Objective 1)
    c. Determining the number of fish populations for each target species within and among Project reservoirs (RSP Objective 2)
    d. Estimating annual N_e for each population of target species. Objective facilitated by using SNP genotypes (RSP Objective 3)
2. Obtain metadata for target species needed to estimate N_e during the ILP study period (RSP Objective 3)
3. As part of #1, collect field data to support base condition genetic analysis of juveniles at Project reservoir tributary spawning grounds (June 9, 2021 Notice Objective 1)
4. Estimate base condition genetic diversity (heterozygosity, within- and among-population variance, and relatedness) for Dolly Varden in Project reservoirs (RSP Objective 1)
5. Expand sample collection and/or coordinate existing samples and activities for out-of-basin and above and below dam analyses. Objective facilitated by using SNP genotypes in combination with coordination of existing information/activities outside of the study area (June 9, 2021 Notice Objective 2)

LP research questions that appear to show congruence with these 2022 objectives include LPs objectives 1-3 above. LP objectives 5-7 might also be addressed by 2022 field collections, should City Light have a shared interest in pursuing analyses beyond objectives specified in the RSP/NOA.

4.2 Study Taxa

Native trout and char within the Project Boundary (*Salvelinus* spp. and *O. mykiss*) constitute age- and genetically structured metapopulations (CFS 2022). A metapopulation is defined as a collection of spatially divided subpopulations that experience a certain degree of gene flow among them (Allendorf et al. 2013). The degree of age and spatial genetic structure varies across taxa; all native trout/char are iteroparous (repeat spawners) and display overlapping generations (have multiple age classes). Currently, the pattern

of metapopulation population structure across the Project reservoirs is not well understood because gene flow among the reservoirs is not yet well described. Hatchery releases derived from Ross Lake brood stocks and hybridization among native and introduced trout/char may also affect the genetics of the native fish. Due to complications of species identification between native species with similar phenotypes and hybridized individuals, target fish taxa for the field study program consist of all *Salvelinus* and *Oncorhynchus* species.

4.3 Sample Collection

The key consideration to sampling for population structure and effective population size is to acquire random samples from the cohorts that comprise the generation(s) of interest within representative subpopulations of the metapopulation (Ryman 2020). Proposed sampling locations for the 2022 field season are listed in Table 1 and depicted on Figure 1. These tributaries are thought to be representative of the subpopulations contributing to the productivity and genetic diversity of native trout/char in the Project Area (in the U.S) based on eDNA surveys and previously identified (during summer 2021) char spawning locations (Ostberg 2022). Sampling adults on spawning grounds is logistically difficult and unethical (e.g., not permitted by regulatory agencies). The proposed sampling approach will therefore target young-of-the-year (YOY) but will include any subadult life stages encountered during surveys. Targeting YOY is less harmful than adults and this early life stage reflects the genetic diversity of adults returning to their natal spawning sites (Garant et al. 2000). Sampling will be conducted using an adjusted version of the recommendations of Whiteley et al. (2012) to avoid overrepresentation of family groups (i.e., Allendorf-Phelps effect [Allendorf and Phelps 1981]). Specifically, three distinct locations (i.e., high, mid, and low elevation) in each tributary will be sampled by collecting fish via backpack electrofishing (following guidelines in *Fisheries Techniques,* Reynolds and Kolz 2012 and according to IACUC permitting conditions) and then excising a small fin clip from up to 50 individuals. A sample size of 50 is expected to provide enough power to detect an effective size between 50 and 500 (i.e., in reference to the "50/500 rule" Franklin 1980) (Waples et al. 2006; Whiteley et al. 2012). However, anecdotal evidence suggests the collection of 50 individuals may be impractical in many tributaries, therefore sample sizes may be scaled by a predetermined survey effort (e.g., initial presence/absence survey, collection of 50 individuals, or up to 90 minutes of electrofishing). Except for fish that appear to be Eastern Brook Trout (*Salvelinus fontinalis*), no effort will be made in the field to collect one taxon over another within each genus (i.e., *Salvelinus* and *Oncorhynchus*) due to phenotypic similarities between species and/or hybrids. This approach is intended to support unbiased inferences about patterns of hybridization across the Project area. Eastern Brook Trout can be present in high densities, and so the maximum sample size of 50 within *Salvelinus* (n=50) could plausibly be reached early in the sampling effort, thus reducing chances that a representative sample of native *Salvelinus* (i.e., Bull Trout or Dolly Varden) genes is achieved. Therefore, only the first 30 apparent Eastern Brook Trout will be sampled, allowing for an additional 20 samples for all other *Salvelinus*. To obtain target sample sizes from each cohort (age class), sampling may need to occur across multiple years, depending on the success of field biologists at encountering individuals from each age class in 2022.

Table 1. Proposed sampling locations for the Reservoir Fish Genetics Study

No.	Drainage	River/Stream Name	Location Description
1	Skagit River, Ross Lake	Hozomeen Creek	Mainstem
2		Freezeout Creek	Mainstem
3		Lightning Creek	Mainstem
4		Three Fools Creek	Mainstem

No.	Drainage	River/Stream Name	Location Description
5		Cinnamon Creek	Mainstem
6		Castle Fork	Mainstem
7		Devils Creek	Mainstem and tributaries
8		North Fork Devils Creek	Mainstem
9		Roland Creek	Mainstem
10		Ruby Creek	Mainstem
11		Canyon Creek	Mainstem up to cascade barrier
12		North Fork Canyon Creek	Mainstem
13		Granite Creek	Mainstem up to cascade barrier
14		Panther Creek	Mainstem up to cascade barrier
15		Pierce Creek	Mainstem
16		Big Beaver Creek	Mainstem and tributaries including Beaver Ponds
17		McMillan Creek	Mainstem
18		Luna Creek	Mainstem
19		Little Beaver Creek	Mainstem above and below barriers
20		Silver Creek	Mainstem
21	Skagit River, Diablo Lake	Thunder Creek	Mainstem
22		Colonial Creek	Mainstem
23	Skagit River, Gorge Lake	Stetattle Creek	Mainstem above and below barrier
24		Pyramid Creek	Mainstem
25		Gorge Creek	Mainstem

4.4 Laboratory Analysis

The goal of laboratory analysis is to genotype extracted DNA from fish samples collected during the 2022 field season. In addition to these new samples, City Light will genotype previously collected samples if they are available (e.g., Pflug et al. 2012; Smith 2010; Ostberg 2022) at the new GT-seq (genotyping-in-thousands by sequencing; Campbell et al. 2015) SNPs to bring the current microsatellite genetic baseline up to date and to obtain higher statistical power for achieving Reservoir Fish Genetics Study objectives. DNA will be extracted from fin clips using Qiagen DNeasy blood and tissue kits following the manufactures protocol. Extracted DNA will be polymerase chain reaction (PCR) amplified and genotyped at hundreds of newly developed GT-Seq-based SNP makers. GT-seq is a form of amplicon sequencing involving a parallel multiplex PCR. GT-seq SNP panels have been developed for Rainbow Trout and Bull Trout and contain diagnostic markers for identifying Dolly Varden and Cutthroat Trout (*O. clarkii*). The fish populations used to develop the SNPs included representatives from the Skagit River Basin, so we expect any possible ascertainment bias to be relatively small (Bohling et al. 2021). For Rainbow Trout, the GT-seq SNP panel includes neutral, taxon-diagnostic, and quantitative genetic markers (e.g., associated with life history) whereas the Bull Trout Panel includes neutral and diagnostic markers. The Bull Trout panel contains 26 markers diagnostic for Eastern Brook Trout and two that can distinguish Dolly Varden. The nearly 220 neutral markers ascertained as polymorphic within Bull Trout were specifically designed to estimate neutral genetic variation, genetic differentiation, genetic assignment tests, and effective population size. For Rainbow Trout, the panel comprises 379 SNP loci (Hargrove et al. 2019) and is used regionally for population genetic analysis. Three of the markers are diagnostic for Cutthroat Trout. Due to lack of

genomic resources for Dolly Varden, including a paucity of published markers that distinguish this taxon from Bull Trout, a targeted analysis, like the approach of Melnik et al. (2020) may be needed to achieve objectives for this species depending on whether they are encountered in the project area (see section 4.5.3).

4.5 Statistical Analysis

4.5.1 Effective size (N_e)

City Light's approach to estimating/indexing effective size (N_e) is based on obtaining the samples and genotypes necessary for estimating the annual effective number of breeders (N_b). N_b is defined as the effective population size of parents that gave rise to a particular cohort in one reproductive cycle. This approach was chosen due to relatively straightforward sampling requirements (genetic collections from fish belonging to single cohorts) and because N_b can be used to produce estimates of generational N_e, which takes overlapping generations and variance in reproductive success into account. Methods for estimating generational N_e are presented in Appendix A if management objectives indicate a need to expand the N_b estimates. To estimate N_b, City Light will use the linkage disequilibrium (LD) method (Waples and Do 2008). Linkage disequilibrium is nonrandom association of alleles among loci. Due to the relatively high number of SNPs on the GT-seq panels (hundreds), City Light will screen for physical linkage prior to estimating effective size to increase the odds that disequilibrium is associated with gametic phase as opposed to physical linkage. The LD method estimates the inbreeding-effective size—or the size of an ideal population (i.e., at Hardy-Weinberg equilibrium) that would result in the same reduction in heterozygosity as in the actual population being considered. When samples are collected from cohorts, as will be the case, the LD method estimates N_b within the year prior to that sampled. The method assumes samples are collected randomly from a well-defined population that does not receive migration and that markers are unlinked, neutral, and in Hardy-Weinberg proportions.

4.5.2 Population structure

Genetic analyses of population structure will be used to address RSP Objectives 1 and 2 and to answer emerging LP questions. Classical analysis of genetic population structure assumes that samples are collected from subpopulations that are known *a priori*. Statistical tests using traditional F-statistics (F_{ST}, F_{IS}, F_{IT}) are then typically used to provide evidence about whether the assumption is valid. Recent analyses of native trout and char in the Project Area using existing microsatellite data suggested there is some structuring associated with individual tributaries (CFS 2022); however, statistical power of the dataset was low (14 to 15 microsatellites). Further, population structuring can be cryptic and weakly associated with individual spawning tributaries. Methods used to estimate the number of genetic populations and to describe population structure will therefore include approaches where subpopulations are predefined (i.e., using F-statistics) and ones where they are undefined or exploratory.

In the approach where subpopulations are assumed to be contained within tributaries, genetic diversity will be measured at each locus and within each sample as the expected heterozygosity (H_S) and the proportion of polymorphic loci. Conformance of genotypic frequencies to Hardy–Weinberg proportions will be assessed using chi-square ($\chi 2$) and/or permutation-based procedures (i.e., exact tests). Departures of observed (H_O) and expected (H_S) multilocus heterozygosity within subpopulation will be assessed using Wright's (1951) F_{IS}. Linkage disequilibrium between all pairs of loci will be assessed using log-likelihood ratio tests. Pairwise genetic differentiation estimated as F_{ST} and significance determined using permutations in the AMOVA (analysis of molecular variance) framework.

In the approach where structure/subpopulations are exploratory, various multivariate and clustering analyses will be used to investigate different genetic groupings. We will use Principal Components

Analysis (PCA) of allele frequencies to parse genetic diversity into orthogonal axes of variation. We will use STRUCTURE to estimate the number of genetically indefinable populations (Pritchard 2000). STRUCTURE infers population structuring by placing individuals into genetic groupings that minimize Hardy-Weinberg and linkage disequilibrium. A range of potential population groupings (*k*) will be explored ranging from panmixia (*k*=1) to complete isolation (*k*= the total number of collections). We will use the Evanno et al. (2005) method to identify the uppermost hierarchical level of structure. This method calculates the largest change in the LnP(D) between each pair of *k* (the number of distinct genetic populations assumed) and *k*-1 for all tests of *k*. Evanno et al. (2005) demonstrated through simulation that Δ*k* (defined as the second order rate of change of the likelihood function with respect to *k*) shows a clear peak at the true value of *k*.

4.5.3 Hybridization

Understanding the amount and pattern of hybridization among taxa in the Project area is germane to City Light objectives (RSP objectives 1,2, and 4), is important to LPs (Question 7), and is crucial to interpretation of genetic data in general. As each genus will be sampled randomly in the field with respect to morphology (e.g., no effort to collect Bull Trout versus Dolly Varden), inferences about hybridization are expected to be relatively unbiased. To estimate the proportion of genetic admixture within individuals, we will calculate a hybrid index based on genotypes at taxon diagnostic loci. The index ranges from 0.00 for the taxon of interest (i.e., Bull Trout [no interspecific alleles]) to 1.00 for the alternative taxon (i.e., Dolly Varden [two interspecific alleles at each locus]) and is calculated by dividing the total number of interspecific alleles in an individual by 2x, where x is the number of diagnostic loci analyzed (Allendorf et al. 2012). F1 hybrids have a hybrid index of 0.50 and are heterozygous for alleles from the parental taxa at all diagnostic loci (i.e., individual's genome is half each species). Subsequent successful reproduction of F1 hybrid individuals with non-hybrids (i.e., a backcross) results in lower magnitudes of inferred hybrid index. Other methods of hybrid analysis may also be used, including Bayesian techniques such as those implemented in program STRUCTURE (Pritchard et al. 2000) or NEWHYBRIDS (Anderson 2003).

The statistical power to correctly identify hybrid individuals is directly related to the number of taxon-diagnostic loci analyzed, the magnitude of hybrid index targeted, and whether inference is being made at the individual or population level. For example, using the 26 diagnostic markers for Eastern Brook Trout on the proposed GT-Seq SNP panel (Bohling et al. 2021), the probability of detecting a multigenerational back cross (5% admixture) in a single fish is 96 percent, but approaches unity (100 percent) at the population level with any appreciable sample size (i.e., n≥10). Yet, given the two diagnostic markers present that distinguish between Bull Trout and Dolly Varden (Bohling et al. 2021), the statistical power to detect the same proportion of admixture in a single fish is 10 percent. Importantly, with a sample size of 30, the power increases to 95 percent. Fortunately, approximately 10 additional markers diagnostic for Dolly Varden not described in Bohling et al. (2021) are available (R. Taylor Personal Communication, April 14, 2022). If no hybrids are detected in collections with reasonable sample sizes (i.e., n≥10; Probability of detection > 0.80) then no further hybrid analyses for Dolly Varden will be considered. However, if hybridization is detected, City Light will genotype all *Salvelinus* at these additional 10 diagnostic markers, which would increase power of detection to approximately 71 percent in any individual fish.

5.0 SUMMARY

This memo describes the field, laboratory, and analytical methods City Light will use in 2022 to meet the objectives of the Reservoir Fish Genetics Study. The methods are focused on describing genetic population

structure and estimating effective size. The described approach will address RSP and NOA objectives and may answer some LP questions by producing key data for: (1) establishing an updated genetic baseline of allele frequencies at newly developed GT-seq SNP markers; (2) estimating diversity within and among subpopulations; (3) providing baseline data for completing above- and below-dam genetic comparisons; (4) comparing genetic relationships of the Project area to other regions/core areas in Washington State, British Columbia, and the Pacific Northwest; (5) enumerating the number of genetically identifiable populations up- and downstream of Gorge Dam; (6) estimating the frequencies of alleles associated with adaptive traits like migration timing; (7) estimating the proportion of admixture with hatchery-origin fish; and (8) describing patterns of hybridization. Since revisions of LP questions are ongoing, the degree to which the current sampling program can address LP questions remains uncertain. Regardless, next steps toward addressing the objectives in this memo include Expert Panel support of the proposed approach and then beginning field logistics planning including procurement of the necessary permits to collect the data of interest. To achieve the objectives and remain on schedule for a field program in 2022, the goal is to have the necessary permits with sufficient time such that sampling can begin in July of 2022.

6.0 LITERATURE CITED

Allendorf, F.W., S. Phelps. 1981. Use of allelic frequencies to describe population structure. Canadian Journal of Fisheries and Aquatic Sciences, 38, 1507–1514.

Allendorf, F. W., G. H. Luikart, and S. N. Aitken. 2013. Conservation and the genetics of populations. Wiley, New York.

Anderson, E. C. (2003). User's guide to the program NewHybrids version 1.1 beta. University of California, Department of Integrative Biology, Berkeley.

Bohling, J., J. Von Bargen, J. M. Louden, M. Small, T. Delomas, R. Kovach. 2021. Developing a Standardized Single Nucleotide Polymorphism Panel for Range wide Genetic Monitoring of Bull Trout. North American Journal of Fisheries Management 41: 1920–1931.

Campbell, N. R., S. A. Harmon, S. R. Narum. 2015. Genotyping-in-thousands by sequencing (GT-seq): a cost-effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources 15(4): 855–67.

Cramer Fish Sciences (CFS). 2022. FA-06 Reservoir Native Fish Genetics Baseline Study – Existing Genetics Data Review. Provided to Seattle City Light, the Expert Panel, and License Participants. Tuesday, January 04, 2022.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14(8):2611-2620.

Franklin, I. R. 1980. Evolutionary change in small populations. In M. Soulé & B. Wilcox (Eds.), Conservation biology: An evolutionary-ecological perspective (pp. 135–149). Sunderland, MA: Sinauer Associates.

Garant, D., Dodson, J. J., & Bernatchez, L. 2000. Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar* L.). Molecular Ecology, 9(5), 615-628.

Hargrove, J. S., T. A. Delomas, J. McCane, M. Davison, M. R. Campbell, R. L. Horn, and S. R. Narum. 2019. Chinook and Steelhead Genotyping for Genetic Stock Identification at Lower Granite Dam. Idaho Department of Fish and Game Report 19-08. Annual Report, BPA Project 2010-026-00.

Jones, O. R., J. Wang. 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10: 551–555.

Laikre, L., F. Olsson, E. Jansson, O. Hössjer, N. Ryman. 2016. Metapopulation effective size and conservation genetic goals for the Fennoscandian Wolf (*Canis lupus*) population. Heredity 117(4): 279–289. https://doi.org/10.1038/hdy.2016.44.

Melnik, N. O., G. N., Markevich, E. B. Taylor, A. V. Loktyushkin, E. V. Esin. 2020. Evidence for divergence between sympatric stone charr and Dolly Varden along unique environmental gradients in Kamchatka. Journal of Zoological Systematics and Evolutionary Research. 58(4):1135-50.

Olsson, F., L. Laikre, O. Hössjer, N. Ryman. 2017. GESP: A computer program for modelling genetic effective population size, inbreeding and divergence in substructured populations. Molecular Ecology Resources 17(6): 1378–1384. https://doi.org/10.1111/1755-0998.12673.

Ostberg, C. 2022. Char eDNA Surveys Progress Report. U.S. Geologic Survey. Personal communication to E. Lowery, Seattle City Light, on March 9, 2022.

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

Reynolds, J.B., and A.L. Kolz. 2012. Electrofishing. Pages 305-361 in A.V. Zale, D.L. Parrish, and T.M. Sutton, editors. Fisheries techniques, 3rd edition. American Fisheries Society, Bethesda, Maryland.

Ryman, N., L. Laikre, O. Hössjer. 2019. Do estimates of contemporary effective population size tell us what we want to know? Molecular Ecology 28: 1904–1918. https://doi.org/10.1111/mec.15027.

Waples, R. S. 2002. Evaluating the effect of stage-specific survivorship on the Ne/N ratio. Molecular Ecology 11(6): 1029–1037.

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

Waples, R. S., and C. Do. 2008. LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753-756.

Waples, R. S., C. Do, and J. Chopelet. 2011. Calculating Ne and Ne/N in age structured populations: a hybrid Felsenstein-Hill approach. Ecology 92: 1513–1522.

Waples, R. S., Antao T, Luikart G. 2014. Effects of overlapping generations on linkage disequilibrium estimates of effective population size. Genetics 197: 769–780.

Whiteley, A. R., J. A. Coombs, M. Hudy, Z. Robinson, K. H. Nislow, B. H. Letcher. 2012. Sampling strategies for estimating brook trout effective population size. Conservation Genetics 13: 625–637.

Whiteley, A. R., J. A. Coombs, M. J. O'Donnell, K. H. Nislow, and B. H. Letcher. 2017. Keeping things local: Subpopulation Nb and Ne in a stream network with partial barriers to fish migration. Evolutionary Applications 10: 348-365.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.

Wright, S. 1951. The genetical structure of populations. Annals of Eugenics 15:323–354.

7.0 APPENDIX A: OPTION TO ESTIMATE GENERATIONAL EFFECTIVE POPULATION SIZE (N_E)

City Light may choose to expand estimates of N_b into estimates of generational N_e. Generational N_e is fundamental to evaluating trends in population viability (or extinction risk) because N_e determines how evolutionary forces act to change the genetic composition of populations (e.g., selection, mutation, gene flow, genetic drift). N_e determines, for example, the amount of diversity a population loses each generation through genetic drift and the rate of allele frequency change through natural selection.

N_e is a demographic parameter that is affected by many factors, such as variance in reproductive success, age structure, and gene flow among (sub)populations. There are numerous ways to calculate N_e using molecular techniques and each measurement can reflect a different type of effective population size (e.g., inbreeding N_e, drift N_e, coalescent N_e, etc.). Ryman et al. (2019) provided mathematical definitions for 18 different calculations of N_e (a subset of definitions pertinent to this memo are presented in the Appendix C); importantly, the metrics are considered equal under the theoretical conditions of an ideal, isolated population of constant size (i.e., a Wright-Fisher population). Yet, as is often the case in nature, the "realized" values of various N_e metrics can differ dramatically in natural systems due to deviation from "ideal" conditions, such as in the presence of spatial and temporal barriers to reproduction (population genetic structure).

To maximize utility of estimated N_e for conservation and management of fish populations within the Project Boundary, the approach for estimating effective size of native trout should consider how population structure and other potential deviations from "ideal conditions" could cause bias in estimated N_e. Some of this information is available from the Reservoir Genetics Study activities in Year 1, and demographic information will be refined from Year 2 collections. Ryman et al. (2019) showed how accounting for the genetic structure within a metapopulation context, with respect to the number of subpopulations, their size, and pattern for connectivity, can support modelling (calibrating) expected magnitudes of various calculated N_e metrics. Accounting for this structure will be key to addressing City Light and LP objectives because of the apparent genetic structure observed among subpopulations of native trout in the Project area (CFS 2022). Additionally, for species with multiple age classes, Waples et al. (2014) showed the effects of overlapping generations and variance in reproductive success on N_e metrics, which can be adjusted using estimates (or expert opinion) of adult life span, age-at-maturity, and fecundity. The technical approach described below will use these published advanced analytics. Further, the sampling design proposed specifically provides for spatially and temporally replicated tissue collections that will allow for bias corrected (adjusted) estimates of the annual form of N_e, the cohort-specific effective number of breeders (N_b). N_b is a steppingstone for estimating the population values over an entire generation (i.e., all age classes; generation). In other words, N_b can be estimated using data collected during the 2022 field season to provide insight into population viability while also establishing the needed foundation for calculating generational N_e (or trend). Importantly, N_b is more easily estimated and interpreted than N_e for organisms such as native trout within the Project that deviate from "ideal" mathematical population modeling, representing non-panmictic metapopulations exhibiting spatial and temporal genetic structure with overlapping generations. A key advantage of N_b is that annual estimates across space and time can be combined with information on genetic structure and demography to provide calibration and a contemporary estimate of generational N_e (Waples et al. 2014; Whiteley et al. 2017).

Returning to the relationship of N_e and conservation science, a reliable estimate of generational N_e (i.e., N_e that measures genetic change from generation t to t + 1) is foundational to conservation planning because it provides the metric used to index population viability under the "50/500 rule" (Franklin 1980). The 50/500 rule is a general rule-of-thumb conservation science has operated under for decades that

states the effective population size should not be less than 50 in the short-term, and not less than 500 in the long-term. The short-term rule is based on well-documented decreases in fitness (reproductive success) due to inbreeding when N_e falls below approximately 50, as animals do not tolerate inbreeding well. The long-term rule is based on the loss of additive genetic variation – literally the effect of adding a specific allele into a population – that is important for potential local adaptation. In the context of a metapopulation (e.g., native trout in the Project area constitute age- and genetically structured metapopulations [CFS 2022]), Laikre et al. (2016) recommended that long-term genetic viability should imply that the rate of inbreeding in the entire metapopulation (N_{eMeta}), as well as in the separate subpopulations (N_{eX}), should be greater than 500 due to the risk of accumulation of inbreeding within subpopulations.

Further technical specification regarding sampling and analysis is described below, but the N_e definitions of interest to the 50/500 rule are the inbreeding effective size (N_{eI}) and the additive genetic variance effective size (N_{eAV}). Due to relative ease of sampling and analysis, N_{eI} and N_{eAV} are commonly estimated using the "linkage disequilibrium" effective size (N_{eLD}) and the "variance" effective size (N_{eV}), respectively. As mentioned, studies have shown that N_e estimated using these metric types are biased in age-structured and subdivided metapopulations (Waples et al. 2014; Ryman et al. 2019). For example, age structure can upwardly or downwardly biased N_e estimates depending on life history and vital rates (Waples et al. 2014). In spatially subdivided metapopulations, both local (subpopulation) and global (metapopulation) estimates can be grossly underestimated. While these issue present obvious challenges to estimating N_e and interpreting those values for native trout within the Project Boundary, the sampling design and analytical approach accommodate these biological realities and provide a foundation for future viability evaluations.

7.1 Analysis of generational N_e using annual estimates of N_b

To estimate generational N_e, the approach developed by Waples et al. (2014) and empirically applied Whiteley et al. (2017) to Brook Trout will be used. Briefly, adjusted subpopulation-specific N_{b-LDNe} will be calculated following the equation $N_{b(Adj)}=N_b/(1.26-0.323x(N_b/N_e))$, where $N_{b(Adj)}$ are bias-adjusted values of raw N_{b-LDNe}. The ratio N_b/N_e will be obtained separately for each subpopulation and taxon using the computer program AgeNe (Waples et al. 2011). N_b will be the subpopulation-specific harmonic means of N_{b-LDNe}. $N_{b(Adj)}$ is then divided by the ratio of N_b/N_e from AgeNe to obtain subpopulation-specific N_e (hereafter $N_{e(Adj)}$). A life history table constructed for each taxon based on demographic data (either newly derived data, obtained from literature, and/or expert best professional judgment) will be used for the AgeNe analysis. AgeNe assumes constant population size and stable age structure (Waples et al., 2013) to obtain the N_b/N_e ratio. AgeNe implements an index of overdispersion of reproductive success of same-age, same-sex individuals termed the Poisson scaling factor (PSF; Waples et al., 2013). A PSF derived by fitting a negative binomial model to full-sibling family size distributions for each cohort will be used by using the program COLONY to estimate family sizes (Jones and Wang 2010). The fitted negative binomial will be used as an estimate of k (mean reproductive success) and the variance used as an estimate of Vk. The ratio Vk/k represents an unscaled PSF, which will be scaled using equation 3 from Waples (2002) to obtain a PSF (Vk2/k2) scaled by the expected k at a constant population size. The mean and variance from the fitted negative binomial distributions will be used as k1 and Vk1, respectively in this equation.

In addition to age structure, Ryman et al. (2019) showed that estimates of N_{eI} based on N_{eLD} can be biased low in the face of population structure and that calibration using apparent patterns of gene flow can support interpretation. Accounting for population structure is crucial to Rainbow Trout in the Project Boundary because of how subpopulations within the Project are genetically structured (CFS 2022) and due to the history of gene flow from hatcheries. Accounting for population structure is also crucial to Bull

Trout due to observed genetic structure among Project Boundary reservoirs. A similar approach to Laikre et al (2016) and Ryman et al. (2019) will be applied to understand how genetic structure could affect estimates of N_e. The program GESP (Olsson et al. 2017) will be used to calibrate estimates of N_e by simulating different metapopulation genetic structure for each taxon. GESP is an R-based computer package designed to model short- and long-term patterns of genetic differentiation and effective population size of subdivided populations. The algorithms performed by GESP allow 1) exact computation of global and local inbreeding and eigenvalue effective population size, 2) predictions of genetic divergence among populations (G_{ST}) as well as departures from random mating (F_{IS}, F_{IT}) while 3) varying (i) subpopulation census and effective sizes, separately or including trend of the global population size, (ii) rate and direction of migration between all pairs of subpopulations, (iii) degree of relatedness and divergence among subpopulations.

8.0 APPENDIX B ASSUMPTIONS AND DATA NEEDS

The estimation of generation N_e follows the methods of Waples et al. (2014) and Whiteley et al. (2017). These approaches involve using the program AgeNe (Waples et al. 2011) to estimate the ratio N_b/N_e, which has required assumptions and data needs. Specifically, AgeNe assumes 1) a stable population that produces a fixed number of individuals that survive at least to age 1, 2) all reproduction occurs at intervals of exactly one time unit, 3) survival and fecundity are independent of events in previous time periods, 4) there is no upper bound to the number of offspring an individual can produce in one breeding cycle, and 5) each newborn has an equal probability of being male or female (Waples et al. 2014).

Data needed to estimate N_b/N_e using AgeNe include an estimate of the number of individuals that survive to age 1 (N_1), mean generation length (Gen), maximum age (ω), and age and sex specific survival and fecundity rates (Waples et al. 2014). These data can be determined from literature values or derived empirically. If sufficient data is available for our target taxa within the study area, AgeNe will be used to estimate N_b/N_e, which we will then use to calculate N_b adjusted for bias due to age structure ($N_{b(Adj)}$) using the first equation in the following table (Table 3 of Waples et al. 2014). However, if insufficient data is available to estimate N_b/N_e for our system, which will likely be the case, the alternative formulae from Table 3 of Waples et al. (2014) will be applied to estimate N_b and N_e, which only require knowledge of two or three traits (adult life span, age at maturity, and coefficient of variation of age-specific fecundity). Using true N_b/N_e from AgeNe, Waples et al. (2014) found that $N_{b(adj)}$ was within 5% of true N_b, while using two traits produced adjusted r^2=0.67 and three traits adjusted r^2=0.84.

Table 2. Formulas to adjust N_e and N_b for biases due to age structure (Table 3 from Waples et al. 2014).

Method	Formula
To estimate N_b	
Using true N_b/N_e	$\hat{N}_{b(Adj)} = \frac{\text{raw } \hat{N}_b}{1.26 - 0.323 \times (N_b/N_e)}$
Using two traits	$\hat{N}_{b(Adj2)} = \frac{\text{raw } \hat{N}_b}{1.103 - 0.245 \times \log(AL/\alpha)}$
Using three traits	$\hat{N}_{b(Adj3)} = \frac{\text{raw } \hat{N}_b}{0.991 - 0.206 \times \log(AL) + 0.256 \times \log(\alpha) + 0.137\ CVf}$
To estimate N_e	
Using true N_b/N_e	$\hat{N}_{e(Adj)} = \frac{\hat{N}_{b(Adj)}}{N_b/N_e}$
Using two traits	$\hat{N}_{e(Adj2)} = \frac{\hat{N}_{b(Adj2)}}{0.485 + 0.758 \times \log(AL/\alpha)}$
Using three traits	$\hat{N}_{e(Adj3)} = \frac{\hat{N}_{b(Adj3)}}{0.833 + 0.637 \times \log(AL) - 0.793 \times \log(\alpha) - 0.423 \times CVf}$

It is additionally important to note that it will be necessary to collect genetic samples from numerous consecutive cohorts in order to reliably estimate N_b. Whiteley et al. (2017) conducted an analysis to approximate the number of cohorts needed to reliably estimate harmonic mean $N_{b\text{-}LDNe}$ for a population when using the method of Waples et al. (2014). For brook trout in their system, they observed that it took 4 or 5 consecutive cohorts for $N_{b\text{-}LDNe}$ to converge to the harmonic mean, and up to 9 cohorts for one subpopulation with higher temporal variation. In light of this finding, we anticipate that generation of reliable estimates of N_b will be dependent upon the ability to collect sufficient (50 per the study plan) genetic samples from the target taxa from at a minimum of 4 or 5 consecutive cohorts.

9.0 APPENDIX C DEFINITIONS

Table 1. Description of notation.

Symbol	Description
N_e	Effective population size per generation that contributes genetically to the next generation.
N_{eMeta}	Total (global) inbreeding effective size of the metapopulation.
N_{eX}	Inbreeding effective size of subpopulation *X*
N_{eI}	Inbreeding effective size, measures the rate of increase in inbreeding
N_{eAV}	Additive genetic variance effective size
N_{eLD}	Linkage disequilibrium effective size
N_{eV}	Variance effective size, measures the rate of gene frequency drift
$N_{e(Adj)}$	Subpopulation-specific estimates of N_e that are calculated by dividing $N_{b(Adj)}$ by the ratio (N_b/N_e)
N_b	Effective number of breeders in one reproductive cycle, important for understanding eco-evolutionary dynamics and mating systems
$N_{b\text{-}LDNe}$	An estimate of N_b based on linkage disequilibrium, as estimated using the program LDNe (Waples and Do 2008)
$N_{b(Adj)}$	An estimate of N_b adjusted to account for bias due to age structure
PSF	Poisson scaling factor, an index of overdispersion of reproductive success of same-age, same-sex individuals
k	Mean reproductive success, mean number of offspring produced per parent per time period
k_1	Mean from the negative binomial model fitted to full-sibling family size distributions for each cohort, as described in Whiteley et al. 2017
V_k	Variance in reproductive success (number of offspring produced) among adults in one time period
V_{k1}	Variance from the negative binomial model fitted to full-sibling family size distributions for each cohort, as described in Whiteley et al. 2017
V_k/k	Index of overdispersion for same-age, same-sex individuals, an unscaled PSF
V_{k2}/k_2	A PSF scaled by the expected value of k at a constant population size, as described in Whiteley et al. 2017, where k_2 is the mean of the scaled age-specific fecundity values (b_x') from AgeNe (see Waples et al. 2011).
N_1	The number of individuals that survive to age 1
Gen	Generation length
α	Age at maturity
ω	Maximum age

AL	Adult lifespan, AL = ω-α+1
CVf	Coefficient of variation of age-specific fecundity

RESERVOIR FISH GENETICS STUDY

ATTACHMENT C

YEAR 2 TECHNICAL MEMORANDUM

Technical Memorandum

Date:	Wednesday, January 25, 2023
Project:	Skagit River Hydroelectric Project
To:	Expert Panel for FA-06 Reservoir Native Fish Baseline Genetics Study
From:	Scott Blankenship and Dan Bingham, Cramer Fish Sciences
Subject:	FA-06 Reservoir Native Fish Genetics Baseline Study – Year 2 Summary for Expert Panel

1.0 INTRODUCTION

The Skagit River Hydroelectric Project (Project) is owned and operated by Seattle City Light (City Light) and is undergoing Federal Energy Regulatory Commission (FERC) relicensing with its current license expiring in 2025. The FA-06 Reservoir Native Fish Genetics Baseline Study (Reservoir Fish Genetics Study) was not required by FERC in its July 16, 2021 Study Plan Determination; however, City Light implemented the Reservoir Fish Genetics Study as proposed in the Revised Study Plan (RSP; City Light 2021) with the agreed upon modifications described in the "Notice of Certain Agreements on Study Plans for the Skagit Relicensing" (June 9, 2021 Notice).[1] The Year 1 objectives outlined in the Reservoir Fish Genetics Study aimed to use genetic data produced by previous studies to better understand the types of inferences that could be drawn about (1) the genetic diversity and population structure of native fish and (2) to identify possible data gaps that might address topics (questions) of interest submitted by licensing participants (LP) to City Light. Year 1 results were summarized and shared with the LPs and Expert Panel (i.e., Year 1 tech memo). Using inferences gleaned from the Year 1 efforts, a plan was developed and executed that included new field collections, requesting data/samples from state and federal agencies, and generation of new genetic data that was intended to provide foundational information on salmonid populations useful to support the relicensing of the Project. Specifically, the Year 2 objectives of the Reservoir Fish Genetics Study were to:

- Expand sample collection and/or coordinate existing samples and activities for out-of-basin and above and below dam analyses.
- Continue data collection to address heterozygosity, within- and among-population variance, and relatedness for Dolly Varden in Project reservoirs.
- Gather additional data needed to estimate N_e for each population of Bull Trout (*Salvelinus confluentus*), Rainbow Trout (*Oncorhynchus mykiss*), and Dolly Varden (*S. malma*).
- Seek input from LPs and advice from an Expert Panel on whether and how genetics information or other monitoring methods can be used to inform future evaluation of reservoir fish abundance, habitat use, and migration timing.

[1] Referred to by FERC in its July 16, 2021 Study Plan Determination as the "updated RSP."

2.0 METHODS

2.0 Sampling

2.0.1 Field Collection

To accurately estimate genetic diversity and effective population size, it is important to acquire genetic samples from randomly selected fish from the cohorts that comprise the generation(s) of interest within subpopulations that are representative of the metapopulation in the study area (Ryman et al. 2019). Likewise, to make comparisons between study area populations and populations occurring outside of the study area (Project vicinity), representative samples from the relevant out-of-basin populations are also needed. City Light attempted to obtain representative samples by acquiring new tissues, requesting previously collected tissues/genotypes from project partners, and by using standardized molecular methods that will facilitate comparisons to future genetic data collection efforts.

New tissues were acquired from within the Project vicinity during the 2022 field season and are listed in Table 1. These tributaries were thought to be representative of the subpopulations contributing to the productivity and genetic diversity of native trout/char in the study area (in the U.S.). Tributaries were chosen based on written requests submitted by LPs during year 1 and were also based on inferences garnered from eDNA surveys and previously identified char spawning locations (Ostberg 2022). The field season presented several challenges, including an active wildfire season, landslides, and other access/safety issues due to the rugged landscape and remote sampling locations. As a result, 14 of the 20 proposed tributaries from Ross Lake were sampled, both tributaries proposed for Diablo Lake were successfully sampled, and 2 of the 3 tributaries for Gorge Lake were sampled. Notably, some tributaries were surveyed but no fish were encountered. The absence of fish may have been associated with natural barriers to their migration, which were measured and determined during the FA-07 Reservoir Tributary Habitat Assessment (as shown in Figure 1).

Backpack electrofishing was used to capture native char and trout for tissue and scale samples. The sampling approach targeted young-of-the-year (YOY) fish to ensure robust N_b estimates[2], but samples also included any subadult or non-spawning life stages encountered during surveys. Targeting YOY is less harmful than adults and this early life stage is often thought to reflect the genetic diversity of adults returning to their natal spawning sites (Garant et al. 2000). Sampling was conducted using an adjusted version of Whiteley et al. (2012) recommendations to avoid overrepresentation of family groups (i.e., Allendorf-Phelps effect; Allendorf and Phelps 1981). Specifically, sampling targeted three distinct locations (i.e., high, mid, and low elevation) in all primary tributaries that flowed directly into a lake and a single site in all secondary tributaries (i.e., ones that flow into a primary). Electrofishing methods followed guidelines in Fisheries Techniques (Murphy and Willis 1996), Reynolds and Kolz (2012), and according to Institutional Animal Care and Use Committee permitting conditions. A small fin clip and scale sample from up to 50 individuals per reach (i.e., within each high, mid, low) were collected. A sample size of 50 is

[2] N_b is the annual effective size or the number of effective breeders in the parental generation (i.e., the generation that gave rise to that cohort). N_b is more easily estimable and interpretable for organisms with age structure, native trout and char. However, generational N_e is more useful because it provides a direct link to a rich body of population genetic theory.

expected to provide enough power to detect an effective size between 50 and 500 (i.e., in reference to the "50/500 rule"; Franklin 1980; Waples 2006; and Whiteley et al. 2012). Due to the possibility of encountering fewer than 50 individuals at a site, sampling was scaled by a predetermined survey effort (e.g., initial presence/absence survey, collection of 50 individuals, or up to 90 minutes of electrofishing).

We sampled randomly within each genus to understand hybridization and because of the possibility for incomplete lineage sorting, as morphological identification of individuals may not accurately reflect their species identity. Incomplete lineage sorting is a process in which two or more lineages of a species diverge but continue to exchange genetic material (i.e., hybridize), resulting in a pattern of shared genetic ancestry among the lineages. This process can lead to the retention of ancestral polymorphisms in descendant lineages, resulting in a pattern of genetic variation that is inconsistent with the species' phylogeny. Thus, with respect to sampling, no effort was made in the field to collect one taxon over another within each genus (i.e., *Salvelinus* and *Oncorhynchus*) due to phenotypic similarities between species and/or inter-species hybrids. This approach was intended to support unbiased inferences about patterns of hybridization within encountered fish across the study area. There was an exception to this random sampling method for Eastern Brook Trout because they can be present in high densities, potentially causing the maximum sample size of 50 to be reached within *Salvelinus* (N=50) early in the sampling effort and reducing the chances of achieving a representative sample of genes from native *Salvelinus* (i.e., Bull Trout or Dolly Varden). Therefore, only the first 30 apparent Eastern Brook Trout were sampled (15 within a reach), allowing for an additional 20 samples for all other *Salvelinus*.

Age information is needed for individual fish to estimate N_b, the annual cohort-specific effective population size. To estimate age, a scale sample was obtained by removing approximately five scales using the edge of a sharp knife to scrape off scales from above the lateral line and posterior of the dorsal fin. The collected scales were transferred to a small square of filter paper and then placed in an individually labeled scale envelope.

2.0.2 Samples and Information from other Entities

Previously collected tissues or genotypes were acquired from USGS and WDFW. Bull Trout samples and genotypes from downstream of the Project area were not available at the time of this study. USGS provided 663 *O. mykiss* tissues sampled between 2018 and 2021 from across ten tributaries within the Project vicinity. WDFW provided genotypes for 30 *O. mykiss* from Gorge Lake and 428 *O. mykiss* from nearby locations below Gorge Dam. WDFW also provided 180 tissues from Dolly Varden collected within the study area between 2019 and 2020 (T. Seamons, WDFW). Seattle City Light provided an additional 32 Bull Trout tissues collected from the Project reservoirs between 2020 and 2022 (J. Fisher, City Light).

2.1 Laboratory Methods

Genomic DNA was extracted from fin tissue using Qiagen DNeasy 96 Kits on the Qiagen Qiacube following manufacturer's recommendations and eluted in 200 µL polymerase chain reaction (PCR)-grade water. Extractions were then concentrated via desiccation and re-elution into 15 µL buffer AE. One well of each 96-well plate remained empty to be processed as a "no-template" control. All *Oncorhynchus* samples were genotyped at a panel of 354 SNPs developed by the Columbia River Inter-Tribal Fish Commission (CRITFC) (Hess et al. 2018) using the

'Genotyping-in-Thousands by Sequencing' method (GTseq; Campbell et al. 2014). The panel consists of 242 presumably neutral loci, 112 loci linked to putative adaptive genetic variation, three species-diagnostic loci to differentiate *O. mykiss*, *O. clarkii*, and their hybrids, and one sex identification locus. All *Salvelinus* samples were genotyped at a panel of 264 SNPs developed by Bohling et al. (2021) using the GTseq method. The panel consists of 235 presumably neutral loci, 20 taxon-diagnostic loci. to differentiate *S. confluentus*, *S. fontinalis* and their hybrids, eight species-diagnostic loci to differentiate *Salvelinus* taxa (*S. alpinus*, *S. fontinalis*, *S. confluentus*, *S. malma*, *S. namaycush*, and *S. leucomaenis*), and one sex identification locus. For *Oncorhynchus* and *Salvelinus* genera, DNA library preparation methods described in Campbell et al. (2014) were followed. Once size-selected, libraries were Qubit-quantified using the Qubit 1x dsDNA HS Assay Kit, normalized to 4 nM and pooled at equal volumes for sequencing on the Illumina MiSeq (MiSeq® Reagent Kit v3 150 cycle). No more than three libraries were pooled and sequenced at a time. Individuals were genotyped with a custom perl script (GTseq_Genotyper.pl; Campbell et al. 2014), and samples were removed if missing data at more than 65 percent of loci.

The Bull Trout GTseq SNP markers were ascertained specifically to analyze genetic variation within Bull Trout, not within Dolly Varden. Thus, there is a possibility for biased inferences about genetic variation within Dolly Varden (i.e., SNP ascertainment bias). Therefore, following identification of Dolly Varden using the taxon-diagnostic markers analyzed in the GTseq panel, samples were genotyped using an eight-locus microsatellites panel described by Melnik et al. (2020): *Ssosl456*, *Sco218*, *Sco205*, *OtsG253*, *Smm3*, *Smm22*, *Smm17*, *Sco204*. PCR was performed with 10 µL reaction volumes and Qiagen PCR components. All thermal cycling was conducted using a Bio-Rad C1000 Touch Thermal Cyclers. All PCR products were visualized by electrophoresis on an ABI 3730 automated capillary sequencer (Applied Biosystems) contracted from UC Davis Veterinary Genetics Laboratory. Fragment size analysis was completed using Geneious bioinformatics software (Biomatters, Inc., San Diego, California) consistent with the knowledge base available for these genetic markers. Samples were analyzed independently by two people to reduce process errors, with discrepancies in genotype results resolved using consensus. Individuals were retained for analysis if their multi-locus genotypes consisted of at least six (of eight) loci.

2.2 Hybridization

Taxon-diagnostic SNPs were the primary method of identifying individuals that were inter-species hybrids. Three diagnostic SNPs present in the GTSeq-379 panel were used to identify putative *O. mykiss*, *O. clarkii*, and their hybrids. A hybridization index was created that counted the number of Cutthroat Trout *(O. clarkii*) diagnostic alleles present within individual genotypes, which ranged from 0 (nonhybrid) up to 6 (contained only Cutthroat Trout alleles). Any individual with a genotype containing one or more Cutthroat Trout allele was omitted from analysis, given the goals of this project were focused on characterizing non-hybridized *O. mykiss* populations within the study area. Investigation of Cutthroat Trout hybridization in the study area is being undertaken by the NMFS, although the distribution of hybrids observed within the study area is reported in this memo.

For identification of Bull Trout within collections of *Salvelinus*, a positive control representing each taxon (Bull Trout, Dolly Varden, and Brook Trout) was processed alongside field samples to verify species identification results. A sample was preliminarily identified as Bull Trout if

homozygous for Bull Trout alleles in the species diagnostic marker *Salv_SNP_013* (TT). Samples with heterozygous genotypes were identified as hybrids and removed from the analysis. A sample was preliminarily identified as Brook Trout if homozygous for Brook Trout alleles in *Salv_SNP_008* (TT). Samples with heterozygous genotypes were identified as hybrids and removed from the analysis. Taxon identity for each sample was secondarily confirmed using the 20 Brook Trout and Bull Trout diagnostic loci. Samples were identified as Brook Trout or Bull Trout if homozygous for their corresponding diagnostic alleles (see Table 2). Samples with heterozygous genotypes at one or more diagnostic locus were identified as hybrids and removed from the analysis. Individuals were identified as Dolly Varden if their diagnostic genotypes matched those of the Dolly Varden positive control.

2.3 Genetic Diversity within Collections

Once hybridized individuals were removed from the dataset, the study team estimated genetic diversity and tested common assumptions within each taxon. The assumption that genotypes conform to Hardy-Weinberg proportions (HWP) within collections representing well-defined populations is a principle stating that given population allele frequencies, genotype frequencies of homozygotes and heterozygotes will reach segregate into binomial proportions after one generation of random mating. The principle posits that genotype frequencies will remain constant if the population is large and experiences no forces altering allele frequencies (i.e., migration, selection, mutation, or nonrandom mating). This principle predicts relationship between the frequency of genetic variants (alleles) within a population and the genotype frequencies formed within individuals that are combinations of alleles present. Tests for HWP were used as a data quality step, as collections that are inconsistent with statistical expectations of the Hardy-Weinberg law could be interpreted as not being from a single population or as evidence of genotyping concerns. As the primary task of this study was to characterize Rainbow Trout, Bull Trout, and Dolly Varden populations within the study area, evaluating consistency of data with HWP is a useful "population" analysis step.

The study team quantified genetic diversity by calculating the expected heterozygosity (H_s) and measured that metric's deviation from the observed heterozygosity (H_o) within collections using Wright's (1949) inbreeding coefficient F_{IS}. Assumptions that genotypes conformed to Hardy-Weinberg proportions (HWP) and gametic (linkage) equilibrium were tested using exact tests. For the Markov chain parameters, 100,000 dememorizations, 100 batches, and 1000 iterations per batch. In some instances, separate collections from the same watershed were pooled to boost sample sizes. Statistical significance was assessed at the α=0.05 level and was corrected for multiple tests using the sequential Bonferroni method (Rice 1989).

2.4 Genetic Diversity among Collections

We used supervised and unsupervised models to test for genetic structure (allele frequency differences) among collections. A supervised model in population genetics is one in which the researcher has prior knowledge of the population structure and can use this information to make predictions about the population. An unsupervised statistical model is one in which the researcher does not have (or does not use) prior knowledge of the population structure and relies on the data to make predictions about the population. For supervised models, subpopulations were assumed to be the same as collections and the subpopulation boundaries were defined by contemporary

watershed boundaries (i.e., the collections from Lightning Creek were treated as a single subpopulation).

For supervised models, we assumed that samples were collected from well-defined subpopulations characterized by the tributaries from which they were collected. We also assumed the subpopulations were part of a metapopulation whose boundaries were defined by contemporary watershed boundaries. We then tested for genetic population structure using a hierarchical approach in which collections were nested within lakes (e.g., Lightning Creek was treated as a subpopulation nested within Ross Lake). Analysis of Molecular Variance (AMOVA) was used to partition genetic variation within individuals (F_{IT}), among individuals within subpopulations (F_{IS}), among subpopulations (F_{ST}), and among groups (i.e., lakes) (F_{CT}).

Migration-drift equilibrium is an implicit assumption of most supervised genetic models (e.g., F_{ST}-based models) and is useful because it supports predictions about the genetic composition of a population. At equilibrium, the rate of gene flow (migration) is equal to the rate of genetic drift (random changes in allele frequencies). This assumption allows predictions about the genetic composition of a population without having to observe it over a long period of time. This is important because if a population is not in equilibrium, then it may be undergoing some form of admixture, selection, or drift, which can have a significant impact on the genetic makeup of the population. By testing for migration-drift equilibrium, it is possible to gain insight into the evolutionary forces that are acting to alter allele frequencies within the population and improve decisions about how to manage the population. To test for equilibrium, we first tested for isolation-by-distance using Mantel tests of pairwise river and genetic distance. Genetic distance was calculated as linear F_{ST} using GenAlEx. Following a significant Mantel test, the assumption of migration-drift equilibrium was assessed using a second Mantel test of the residuals from the initial fitted line against pairwise geographic distance. At equilibrium, scatter (residuals) should increase with increasing geographical separation, as stochastic drift becomes the dominant force over gene flow shaping allele frequencies at greater distances due to increasing reproductive disconnection. Equilibrium conditions between gene flow and stochastic drift are common among populations of salmonids and reflect the balancing of loss of alleles due to drift against their replacement via gene flow, especially among neighboring populations (Hutchison and Templeton 1999). Yet, due to the complex history of extirpation/recolonization during the Pleistocene glaciations, nonequilibrium/dynamic equilibrium conditions may be expected in recently deglaciated watersheds (Taylor et al. 2003).

For the unsupervised analyses, we used Principal Coordinates Analysis (PCoA), Principal Components Analysis (PCA), and Discriminant Analysis of Principal Components (DAPC). Unsupervised analyses can provide an alternative perspective that are free from *a priori* assumptions about how populations are structured (e.g., the F_{ST}-based approach above assumes structure is centered on tributaries). DAPC analyzes PCs to maximize genetic variance among groups of individuals to support identification of genetic groupings or clusters. DAPC is a Bayesian approach and uses information from all genetic markers to create new axes and then projects the data in a way that maximizes separation of genetic groupings. To achieve this, DAPC uses sequential k-means and Bayesian Information Criterion (BIC) model selection to infer genetic clusters (i.e., the prior). K-means clustering is a method of vector quantization, originally from signal processing, that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype (i.e., prior) of the cluster.

Membership probabilities are then estimated using retained discriminant functions (i.e., posterior probabilities of assignment). DAPC is often preferred over a STRUCTURE-like approach because it does not assume panmixia within genetic clusters and can accommodate more complex structures such as hierarchical or steppingstone (Jombart and Ahmed 2011). We used the "a-score" to avoid overfitting the models, which is simply the difference between the proportion of successful reassignment of the analysis (observed discrimination) and values obtained using random groups (random discrimination) (Jombart 2011). It can be seen as the proportion of successful reassignment corrected for the number of retained PCs.

2.5 Effective Population Size

The effective population size (N_e) is among the most important parameters in conservation because it influences the efficiency of natural selection, gene flow (migration), inbreeding, and loss of genetic and/or trait variation. Monitoring N_e and the annual effective number of breeders (N_b) can facilitate early detection of population declines. The program LDNE (Waples and Do 2008) was used to estimate N_e and N_b in the study area. The software uses LD (correlation of allele frequencies among loci) to estimate the inbreeding-effective size, or the size of an ideal population (i.e., at HWE) that would result in the same reduction in heterozygosity as in the actual population being considered. When the method is applied to samples from individual cohorts, LDNE estimates N_b within the year prior to that sampled. The program assumes samples are collected randomly from a well-defined population that does not receive migration and that markers are unlinked, neutral, and in Hardy-Weinberg proportions. To provide a balance between precision and bias, we used an allele frequency cutoff of 0.02. The bias correction of Waples et al. (2014) for overlapping generations was applied within *Salvelinus*, which is based on estimates of adult life span and age of first reproduction obtained from the peer-reviewed scientific literature.

2.5.1 Identification of cohorts to estimate N_b

The age of each sample was estimated to inform N_b calculations in the study area. Scale collection, preparation, and aging followed the general protocols of Love (2016) and Copeland et al. (2018). Samples were dried and the three most legible scales per fish were mounted between two microscope slides for imaging. High-resolution scale images were captured using a microscope, a digital microscope camera and Image Pro Premier version 9.2 software. A subset of individual scales selected based on size class from each genus were visually analyzed for age based on the number of annuli. Un-aged scales were assigned ages using a length-at-age model described below. Salmonid scales are difficult to age due to their variable life history and individual differences in scale reabsorption during stressful periods, leading to misidentification of annuli (Hernandez et al. 2014). Independent age estimates were performed by at least two reviewers. Age determinations were compared between reviewers; if an age difference occurred between the reviewers, a more senior reviewer resolved the difference. Due to the uncertainty in aging, all age classes are reported as X+ indicating that the age is at least that old, but the exact age is undetermined. A length-at-age key based on the subset of scale aged fish and their fork length was constructed for each genus using the R package {FSA} (Ogle et al. 2022). Fish of unknown ages were assigned to age using the semi-random method (Isermann and Knight 2005) based on the length-at-age key. Fish with lengths outside of the range of the length-at-age key fish were not assigned an age.

2.5.2 Haplotype Diversity Associated with Life History

Potentially adaptive (non-neutral) markers are included in the SNP panel used for genotyping *Oncorhynchus* samples in this study. The panel includes markers for genetic regions indicative of juvenile emigration propensity (frequency of OMY5), adult return timing to freshwater (frequency of OMY28), and adaptive diversity associated with climate, land cover, stream temperature, elevation, wind velocity, solar radiation, and stream network variables. The RSP did not specify any analysis needs regarding adaptive diversity, so these data will remain largely unevaluated. However, a cursory evaluation regarding adaptive diversity associated with juvenile propensity to emigrate (chromosome 5 loci; OMY5) was conducted to demonstrate use of adaptive data generated in this study. No life history information is available for Project *O. mykiss*, so specific blocks of information from OMY5 region (haplotypes) cannot be associated with any fish's behavior; what can be accomplished is to document the different haplotypes observed in the study area and their base condition frequencies. Haplotypes present in the study area and their frequencies were determined using {haplo.stats} package in R (Sinnwell et al. 2022). Global frequencies across all individuals were estimated as well as frequencies within each population.

3.0 RESULTS

3.0 Hybridization

3.0.1 Native Char (*Salvelinus* spp.)

Using 22 diagnostic GTseq SNPs, and assuming hybrids with Brook Trout were successfully removed, the probability of failing to detect 10 percent admixture into Bull Trout in each sample was less than 0.01, suggesting high resolution of the panel to distinguish among *Salvelinus* taxa. In the total collection size of 374 *Salvelinus*, we genetically identified 66 Bull Trout (18 percent), 229 Dolly Varden (61 percent), 47 Brook Trout (13 percent), 24 Dolly Varden x Bull Trout hybrids (6 percent), and eight Dolly Varden x Brook Trout hybrids (2 percent). We did not detect any Bull Trout x Brook Trout hybrids (Figure 2). Hybrids were somewhat widely distributed. Specifically, within Ross Lake tributaries, Dolly Varden x Bull Trout hybrids were distributed across 12 sites including Big Beaver (1), Canyon (1), Granite (1), Hozomeen (1), Lightning (2), Roland (2), Ruby (5), Silver (5). Four hybrids were detected in Thunder Creek (4) (Diablo Lake tributary), and one hybrid detected at-large from Diablo Lake (J. Fisher City Light). One hybrid was detected from Stetattle (1) (Gorge Lake tributary). For Dolly Varden x Brook Trout, hybrids were detected across five sites including Ross lake tributaries, Pierce (1) and Silver (2) creeks, and Diablo lake tributaries Colonial (1),and Thunder (3) creeks, plus one in a collection from the Gorge reservoir (J. Fisher, City Light). Figure 3 shows a scatterplot of the first two PCs for all *Salvelinus* estimated from all 263 GTseq SNP markers and shows clear distinction among Bull Trout, Dolly Varden, and Brook Trout. Representative photos for each taxon and their hybrids are presented in Attachment F.

3.0.2 *Oncorhynchus species*

Genotypes for 1,425 *Oncorhynchu*s individuals from the study area were assessed for hybridization status using three taxon-diagnostic markers. A hybrid is an individual with parents that were different species or possessed Cutthroat and Rainbow Trout genes. Eighty-seven percent

of individuals collected possessed only *Oncorhynchus mykiss* alleles at all three diagnostic loci. Meaning, given the power to detect individuals of mixed ancestry afforded by three loci, non-hybrid *O. mykiss* represented 87 percent of *Oncorhynchu*s collected randomly within tributaries. Of the 13 percent hybridized individuals observed (N=190), the hybrid index ranged from 1-6. For this study, an individual with any Cutthroat allele was considered a hybrid and was omitted from the dataset used for genetic analysis. The study team wanted to be conservative and exclude all hybrids. Observing hybrids means there is ongoing reproduction among Cutthroat and Rainbow Trout present in the study area. The geographic distribution of hybridized individuals observed is shown in (Figure 4). More hybridized individuals were observed within tributaries from the southern portion of Ross Lake.

3.1 Genetic diversity within collections

3.1.1 Native Char (*Salvelinus* spp.)

The 66 Bull Trout collected were from Big Beaver (N=2), Colonial (N=1), Granite (N=1), Ruby (N=12), Stetattle (N=13), Thunder (N=2), the mouth of Lightning Creek (2), Gorge Reservoir (10), Ross Reservoir (22), and mainstem Skagit River downstream of the Project (N=1). Thus, sample sizes within individual collections were too small to provide robust estimates of genetic diversity using the supervised models of population structure. Pooling samples from individual tributaries into reservoir-based groups resulted in collections of size 3 in Diablo Lake, 23 in Gorge Lake, and 39 in Ross Lake (Table 3). From the initial suite of 235 neutral markers, 200 were removed from the analysis due to lack of polymorphism (i.e., Minor Allele Frequency less than [MAF] 0.01). When tested within the pooled collections (Ross, Diablo, Gorge), five markers showed significant deviations from HWP at the α=0.05 level with four tests being significant following sequential Bonferroni correction. Two of the markers (*ScoRAD6812* and *ScoRAD4566*) showed substantial heterozygote excess in Ross Lake (F_{IS}=-0.55 and -0.52, respectively) and Gorge Lake (F_{IS}=-0.89 and -0.60, respectively) and were therefore removed from the analysis. The final Bull Trout dataset contained 33 GTseq SNP markers. Fifty-six out of 1,560 (~4 percent) pairwise tests for LD were significant at α=0.05 level and nine were significant following sequential Bonferroni correction (LD was not estimated in Diablo due to small sample size). Gene diversity (H_S) at the 33 markers was 0.29 (SD=0.16) in Ross Lake and 0.33 (SD=0.15) in Gorge Lake. F_{IS} was 0.03 in Ross Lake and -0.03 in Gorge Lake.

The 229 Dolly Varden sampled during the 2022 field season were combined with 210 Dolly Varden shared by WDFW during 2019 and 2020 for a total of 439. However, 25 individuals were removed from the analysis due to missing genotypes at two or more loci. One additional individual was removed due to an identical genotype in another individual (matching individuals 20NW0447 and 20NW0453 from Lightning Creek, 20NW0447 was retained). The final dataset contained 413 Dolly Varden from 13 tributaries (Table 4). A total of 102 alleles across the eight microsatellites were observed. Nine of 48 exact tests showed significant deviations from HWP at the α=0.05 level and two were significant after sequential Bonferroni correction. All eight loci were retained because none of the markers consistently deviated from HWP across collections. Mean H_S was 0.65 (SD=0.05) and was lowest in Lightning Creek 0.58 (0.28) and highest in Colonial 0.72 (SD=0.21). Mean F_{IS} across collections was 0.00 (SD=0.04) and ranged from -0.05 in Colonial to 0.04 in Lightning and Ruby. Eighteen of 308 (5 percent) pairwise tests for LD were significant at the α=0.05 level, but none were significant following Bonferroni correction for multiple tests.

3.1.2 Rainbow Trout (*Oncorhynchus mykiss*)

After removal of 190 hybridized individuals and one individual for not having a verified collection location, 1,234 *O. mykiss* were retained for analysis (Table 5). The dataset was screened for neutral genetic loci that were uninformative (i.e., MAF of less than or equal to 0.01). Seven loci matched this criterion and were removed from dataset (Omy104569114, OmyG3PD2371, Omyb9164, Omycarban1264, Omycyp17153, Omygadd45332, and Omysys1188), resulting in 235 neutral loci (per individual) retained for genotypes.

Prior to conducting genetic analysis on populations, the populations to analyze must be determined. A heuristic assessment of coherent genetic groups was conducted using DAPC. An initial exploratory DAPC used 200 genetic PCs and number of clusters (k) from 1-15, which considered 98.3 percent of observed genetic variance in the dataset. The initial DAPC was evaluated further, as importantly, retention of large numbers of PCs with respect to the number of individuals analyzed can over-fit the discriminant functions. If this occurs, individual membership in selected k clusters can become statistically unreliable, as discriminant functions could become flexible enough to discriminate any number of clusters, overinflating best-fitting clusters. The trade-off between power of discrimination and over-fitting can be measured by the a-score.

Implementation of the a-score procedure repeated DAPC on the dataset using from 1 up to 50 PCs sequentially, with seven PCs estimated to optimize the proportion of successful reassignment corrected for the number of retained PCs (data not shown). DAPC was then rerun on the *O. mykiss* dataset using seven retained PCs (instead of the initial 200), which considered 24.4 percent of the observed genetic variance in the dataset. When considering BIC-based selection of various possible number of clusters (k), the primary inflection point was for k=4 (i.e., four genetic clusters data not shown), which provided a data driven starting point for the potential number of populations present in the study area. At the risk of causing confusion, there were a series of preliminary genetic analyses conducted on iterations of clustering *O. mykiss* individuals that will not be detailed here. Instead, the logic and reasoning will be described in brief on how both k and classification of individuals to populations was achieved using the observed data.

While individual probabilities for cluster membership were statistically reliable at k=4, a majority of fish resided in a single genetic cluster that was inconsistent with this cluster representing a single population given subsequent genetic analysis (e.g., HWE). At k=5, that same large genetic cluster split into two genetic clusters (Figure 5, cluster 3 and 5). Itemization of membership probability values for each individual *O. mykiss* are not shown. Yet, membership probabilities for all fish analyzed showed that individuals from the three smaller genetic clusters were distinctive (clusters 1, 2, and 4), and individuals attributed to cluster 3 and 5 had varying probability. Cluster 1 was Little Beaver Creek, cluster 2 was Pyramid Creek, cluster 4 was Three Fools Creek, with all remaining Project *O. mykiss* residing in either clusters 3 or 5 (Table 6). At k=6, membership probabilities did not improve classification of individuals (data not shown), so k=5 was determined to be the logical categorization based on discriminant analysis of genotypes. At k=5, genetic clusters 3 and 5 were still inconsistent with these clusters representing single populations (data not shown). Therefore, the genetic dataset was partitioned by both DAPC cluster and geographic location for subsequent population analysis to maximize consideration of genetic variation observed.

HWE was estimated for each genetic locus (235 loci) within each population (28 populations). In this data configuration, a majority of loci conformed to HWE expectations, and no locus failed HWE across on populations (data not shown), indicating the genetic loci were suitable for population analysis of study area *O. mykiss*. Relatedness among individuals within each population was estimated using Rxy metric. Mean Rxy was negative for all populations except Lightning and Three Fools Creeks, with confidence intervals overlapping zero (data not shown). The small number of individuals observed with Rxy greater than 0.5 were not omitted from data analysis. Genetic diversity (both observed heterozygosity and expected heterozygosity under Hardy-Weinberg model) is shown in Figure 6 for populations with greater than 15 samples. Diversity is highest in Stetattle Creek and lowest in Three Fools Creek, with expected heterozygosity higher than observed heterozygosity. This distribution of diversity resulted in positive F_{IS} values for many populations (Table 6), meaning there was a reduction in heterozygosity observed from what was expected under assumptions of Hardy-Weinberg model. Another measure of genetic diversity, allelic richness, was highest in Stetattle Creek (1.321) and lowest in Three Fools Creek (1.116).

3.2 Genetic diversity among collections

3.2.1 Native Char (*Salvelinus*)

For supervised analyses, Bull Trout were grouped by their reservoir of origin due to small sample sizes within individual tributaries (mean sample size within tributaries = 7). AMOVA based on reservoir groupings showed that reservoirs account for 5 percent of genetic variation (F_{ST}=0.05; P<0.01) and that variation within individuals accounts for the remaining 95 percent. Mean pairwise F_{ST} between reservoirs was F_{ST} = 0.03 (95 percent CI: 0.02 to 0.05) (Table 7). The highest divergence occurred between Diablo and Gorge lakes (F_{ST}=0.05) and the lowest between Gorge and Ross lakes (F_{ST}=0.03). By contrast, sample sizes tended to be large enough within Dolly Varden to analyze them by tributary (mean sample size = 32). Mean pairwise F_{ST} between collections of Dolly Varden was F_{ST}=0.05 (95 percent CI: 0.037, 0.055). The highest divergence occurred between Lightning Creek and Roland Creek (F_{ST}=0.16). The lowest divergence occurred between Silver and Stetattle Creek (F_{ST}=-0.03), a negative result likely due to small sample size (n=6 and 7, respectively). The next-lowest divergence occurred between Ruby Creek and Canyon Creek (F_{ST}=0.003), which was not unexpected, given Canyon and Ruby Creeks are only nominally distinct (Canyon Creek becomes Ruby Creek in the lower watershed). Unlike Bull Trout, hierarchical AMOVA based on reservoir groupings (i.e., nesting tributaries within reservoirs) did not explain a significant amount of genetic divergence in Dolly Varden (F_{CT}=0.006; P=0.32).

Mantel tests showed a strong relationship between geographic and genetic distance (P<0.01; R^2=0.40), suggesting genetic structure of Dolly Varden is characterized by isolation-by-distance (Figure 7). A positive relationship between geographic distance and the scatter of residuals from the isolation-by-distance analysis was also observed; however, a Mantel test revealed the relationship was not significant (P=0.11, R^2=0.10) and so the null hypothesis that Dolly Varden in the study area are not at migration-drift equilibrium could not be rejected (Figure 8). Sample sizes were not big enough to reliably study isolation-by-distance in Bull Trout.

For the unsupervised analyses, a scatterplot of the first two PCs (37.95 percent) appeared to show genetic structuring (Figure 9) of Bull Trout in the study area, yet specific patterns were visually obscure (i.e., not obviously associated with contemporary watershed boundaries). DAPC was implemented to identify and describe clusters of genetically related individuals and summary

statistics of the analysis are presented in the Attachment A. The most optimal *k*-means-based clustering solution occurred when *k*=3 (BIC=98.43), but we also explored *k*=2 (BIC=99.63) and *k*=4 (BIC=98.51) due to comparable model support (Attachment Table B1- B3). DAPC of these prior inferred k-means based clusters using the first 20 principal components explained 94.5 percent of the variance for *k*=2 and *k*=3 (1 and 2 discriminant functions, respectively) and 86.1 percent of the variance for *k*=4 (3 discriminant functions). Scatterplots of the discriminant functions clearly distinguished the inferred genetic clusters visually with little to no overlap among clusters (Figure 10; Attachment Figure C1- C3). The posterior probability of assignment back to the prior inferred clusters was 1.00 for all *k*, suggesting clear-cut genetic groups exist in the study area. Nevertheless, the a-score (an index of overfitting) suggested that 20 principal components was likely an overfit of the data and so the discriminant analysis was rerun for *k*=3 (i.e., the most supported model) using a more optimal number of 6 PCs (Attachment Figure C5). Posterior assignments were not associated with any apparent contemporary physical features, such as reservoirs or tributaries (Attachment Figure D1). The composition plot of posterior assignments for *k*=3 did not consistently group Bull Trout into collections based on current reservoir boundaries, suggesting contemporary reservoirs may not provide a complete picture of the genetic structure of Bull Trout in the study area (Figure 11). Specifically, the most supported model placed Bull Trout into three genetic clusters that were well mixed between Ross Lake and Gorge Lake.

The PCoA clearly distinguished three genetic groupings of Dolly Varden that did not obviously/visually reflect any temporal (i.e., age-related) or spatial (i.e., distance/watershed) structure (Figure 12 top panel). That is, the groups were not comprised of distinct cohorts or fish from specific watersheds. Rather, each group contained Dolly Varden from all field seasons (i.e., WDFW 2019 and 2020 plus CFS 2022), ages, and nearly all watersheds. Yet within each group, structuring was consistent with contemporary watershed boundaries. Thus, collections from any given watershed simultaneously (paradoxically) clustered together within a group, yet separately across groups. This pattern stood out relative to the isolation-by-distance analysis, which on its own, suggested that closer watersheds (i.e., neighbors) should consistently contain Dolly Varden populations with similar allele frequencies. Such a discrepancy could be created by cryptic genetic structure associated with admixture (i.e., hybridization) or hierarchy (e.g., hierarchical island model). This observation was examined in further detail below.

For Dolly Varden, the first two PCs showed genetic structuring, and like Bull Trout, the specific patterns or drivers were not visually obvious (Figure 13). This stood in contrast to the PCoA of Dolly Varden, which clearly distinguished three genetic groupings. Considering all 101 PCs in the DAPC, the most optimal k-means-based clustering solution for Dolly Varden occurred when *k*=12 (BIC=339.01). Discriminant analysis of these 12 clusters using the first 40 principal components and 11 discriminant functions explained 94 percent of the variance for *k*=12. Visually, a scatterplot of the first two linear functions showed substantial overlap among the inferred genetic clusters (Figure 14; Attachment Figure C4). Nevertheless, the posterior probability of assignment of individuals back to the 12 inferred clusters was very high (95 percent accurate), suggesting the 12 groups reflect a tangible and substantive underlying genetic structure in Dolly Varden. Nevertheless, the a-score suggested 40 principal components likely provides an overfit of the model and so the discriminant analysis was re-run for *k*=12 (i.e., the most supported model) using a more optimal number of 15 PCs (Attachment Figure C6. The optimized model performed nearly as well, providing 94 percent accuracy of posterior assignments back to inferred clusters. Intriguingly, composition plots of the posterior assignments to the 12 genetic clusters consistently

grouped individuals from very distal watersheds together, again highlighting a contradictory pattern relative to the isolation-by-distance analysis that suggested neighboring populations should consistently contain relatively similar allele frequencies (Attachment E).

3.2.1.1 Cryptic genetic structure within Dolly Varden

As mentioned above, the observation that genetic structure was inconsistent with contemporary watershed boundaries despite strong isolation-by-distance was dissimilar to patterns seen among Dolly Varden from watersheds outside of the Skagit River (e.g., Harris et al. 2015; Taylor et al. 2015). Specifically, posterior assignments using DAPC (k=12) consistently grouped Dolly Varden together that were collected from very distal watersheds and reservoirs (see Attachment E [e.g., Dolly Varden from Lightning Creek clustering with Dolly Varden from Thunder Creek]). Under isolation-by-distance, distal sites are expected to harbor Dolly Varden with dissimilar allele frequencies because the effect of drift is stronger than gene flow over longer distances at equilibrium. For instance, Harris et al. (2015) observed a pattern of isolation by distance in Northern Dolly Varden from the western Canadian Arctic, indicating gene flow occurs primarily among neighboring populations and that gene flow and genetic drift have reached equilibrium following the last glaciation. In the study area, it was therefore suspected that cryptic structure reflecting genetic admixture between divergent groups (i.e., hybridization) or some other form of hierarchy (i.e., hierarchical island/stepping stone) could potentially create a pattern of isolation by distance but that any given fish might have closer ancestry to fish from a distal population by virtue of admixture.

To test the hypothesis that admixture (hybridization) between two unknown populations drives cryptic genetic structure in the study area, the DAPC was forced to acknowledge just two populations, k=2 (recall, the most supported model suggested k=12). In the absence of cryptic admixture (i.e., under the null hypothesis), gene flow among nearby sites is expected to keep allele frequencies relatively similar under the observed pattern of isolation by distance, and so the two groups should simply consist of two groups of "neighbors" (e.g., one group of "northern" sites and one of "southern" sites, hypothetically). Yet, if two divergent unknown populations were present, then proximate sites might not be expected to cluster because the overarching structure reflects the spatial distribution of hybrids and not equilibrium between gene flow and drift. The test revealed a striking genetic divergence distributed across the entire study area (i.e., no geographical or temporal pattern), lending evidence against the null hypothesis of equilibrium and in favor of admixture (Figure 15). Recall, the residual test also suggested nonequilibrium conditions. When the PCoA was updated with the k=2 posterior assignments, structure based on a hierarchy of admixture and watershed boundaries was clear (Figure 12 bottom panel). Ancient hybridization with Bull Trout cannot be ruled out as a driver and samples are being sent to the University of British Columbia to confirm their identity.

3.2.1.2 Scale Age Determination and Length-at-Age Key Assignment for Estimating N_b

Scales were collected to estimate ages of fish in support of estimating effective population size. During the 2022 field season, scales were collected from 255 *Salvelinus* individuals of which 110 were aged. The 110 *Salvelinus* samples contained four age classes (0+, 1+, 2+, 3+, and 5+) with a range of 1-60 individuals per age class (Table 8). Age class 0+ and 1+ displayed the most overlap of fork lengths between age classes while other age classes displayed little overlap (Figure 16).

A total of 201 un-aged *Salvelinus* samples were assigned ages based on the length-at-age key for a total of 311 individuals and had a range of 2-175 individuals per age class (Table 9). Age 0+ fork lengths ranged from 60–118 mm, age 1+ fork lengths ranged from 76–188 mm, age 2+ fork lengths ranged from 160–256 mm, age 3+ fork lengths ranged from 323-426 mm, and age 5+ had two individuals with fork lengths 420 mm. There was considerable overlap of fork lengths between age classes 0+, and 1+ (Figure 17).

3.2.1.3 Effective Population Size

Tables 10 and 11 contain summaries of effective size estimates for Bull Trout and Dolly Varden. N_e for Bull Trout was 31.40 (95 percent CI 17.50, 68.80) when all 65 individuals were analyzed as a collection from a single population (i.e., uncorrected mixed-cohort N_e). When samples were divided into groups based on their sampling location (i.e., Ross, Diablo, or Gorge Lake), N_e was 30.9 (95 percent CI 17.3, 90.6) in Ross Lake, 10.9 (95 percent CI 6.1, 22.4) in Gorge Lake, and was inestimable (-1.2) in Diablo Lake due to small sample size (n=3). When samples were divided into the three inferred genetic clusters identified by the DAPC, the effective sizes were 98.50 (95 percent CI 26.20, infinite) for inferred cluster 1 (k1), 6.7 (95 percent CI 2.60, 27.0) for inferred cluster 2, and 24.50 (95 percent CI 12, 173) for inferred cluster 3. The effective number of breeders (N_b) was also attempted to be estimated by grouping individuals into cohorts. Sample sizes within individual cohorts were too small, however, so fish were grouped into two mixed cohort groups: one group consisted of age-0 to 1+ (n=29) and the second group consisted of 2+ (n=36). Raw N_b for age 0 to 1+ was 6.4 (95 percent CI 3.2, 10.4) and the N_b for age 2+ was 47.2 (95 percent CI 23.4, 194). These raw N_b estimates were used to calculate adjusted N_e using Waples et al. (2014) "two trait" correction formula, which produced corrected estimates 12.08 and 88.35 for group 1 and group 2, respectively (Table 10).

For Dolly Varden, uncorrected, mixed cohort N_e was 24.53 (harmonic mean [95 percent CI 18.19, 34.36) and ranged from 20.7 (95 percent CI 13.8, 34.6) in Colonial Creek to 30.5 (95 percent CI 23.4, 41.3) in Big Beaver Creek. When collections were divided into the 12 inferred clusters based on DAPC, harmonic mean N_e was 27.00 (95 percent CI 19.52, 41.10) and ranged from 17.80 (95 percent CI 12.0, 28.2) in cluster 4 to 36.20 (95 percent CI 25.9, 54.7) in cluster 7. Adjusted N_b was 20.61 (harmonic mean [95 percent CI 14.47, 31.24) and ranged from 13.02 (95 percent CI 8.94, 19.90) in Lightning Creek in 2021 to 30.42 in Thunder Creek in 2021. Adjusted N_e was 42.93 (harmonic mean [95 percent CI 26.87, 76.20]) and ranged from 27.24 (95 percent CI 21.02, 45.97) in Lightning Creek in 2021 to 63.13 (95 percent CI 55.24, 91.33) in Thunder Creek in 2021. Following the a posteriori discovery that two highly divergent genetic populations of Dolly Varden could be present in the study area, it was noted that LD due to genetic admixture could produce a downward bias in estimates of N_e. Therefore, N_e was recalculated considering hierarchical structure by dividing collections from each tributary into two groups reflecting the two populations. N_e and N_b estimates approximately doubled (see Table 11).

3.2.2 Rainbow Trout (*Oncorhynchus mykiss*)

3.2.2.1 Among Population Diversity

Pairwise estimates of F_{ST} were calculated amongst populations with greater than 15 samples (Figure 18). These measures can be interpreted as a genetic distance. Little Beaver, Three Fools, and Pyramid creeks were the most divergent Project populations, corroborating the DAPC analysis; however, all pairwise estimates of F_{ST} except one (Granite-1 versus Canyon-1) were

statistically significant (i.e., non-zero). There were some opportunities to compare populations within the same connected tributary. Comparisons were possible for 1) Lightning and Three Fools creeks and 2) Ruby, Canyon, Granite, and Panther creeks. As mentioned, Three Fools Creek was distinctive, so was divergent from the downstream population in Lightning Creek. Ruby, Canyon, and Granite populations from genetic cluster 3 had the lowest F_{ST} values observed. Panther Creek genetic cluster 5 was divergent from other populations in this tributary. Canyon Creek genetic cluster 5 was also more divergent from Ruby, Canyon, and Granite genetic cluster 3 then this population was from Granite Creek cluster 5. Additionally, Ruby, Canyon, and Granite genetic cluster 3 was more similar to adjacent tributaries (e.g., Big Beaver) than to genetic cluster 5 within the same tributary. The global underlying distance pattern observed amongst comparisons between genetic clusters 3 and 5 was that F_{ST} were smaller between cluster 3 populations, irrespective of location, than between cluster 3 and cluster 5 populations. In contrast, while Canyon-5 and Granite-5 exhibited a small F_{ST}, comparisons between cluster 5 populations (Panther-5, Canyon-5 and Granite-5, Stetattle-5) tended to be large.

3.2.2.2 Scale Age Determination and Length-at-Age Key Assignment for Estimating N_b

In relation to estimating the effective population size, age information is necessary to partition sampled fish into single-aged cohort. During the 2022 field season, scales were collected from 407 *Oncorhynchus* individuals, of which 84 were aged. The 84 *Oncorhynchus* samples contained six age classes (0+, 1+, 2+, 3+, 4+, and 5+) with most individuals consisting of younger (age 2 or younger) age classes (data not shown). There was considerable overlap of fork lengths between age class 0+ and 1+ and 1+ and 2+ (data not shown). Less fork length overlap of was observed in older fish, but sample size was smaller. A total of 484 un-aged *Oncorhynchus* samples were assigned ages based on the length-at-age key for a total of 508 individuals (Table 12). There was considerable overlap of fork lengths between adjacent ages for classes 0+, 1+, and 2+; assignment of larger fish was difficult due to few aged fish above 200 mm (Figure 19).

3.2.2.3 Effective Population Size

There are various ways that population by age data can be parsed for estimation of effective population size or annual effective number of breeders (N_b). Given the population analysis above, genetic clusters 3 and 5 would be analyzed separately for each single age cohort. However, the quantity of samples collected and aged from 2022 were insufficient to achieve this configuration. Alternatively, genetic cluster 3, excluding Lightning Creek, had pairwise F_{ST} of approximately 0.02, which is a theoretical threshold for genetic drift connectivity. Therefore, for the initial calculations estimating of annual values for N_b of Project *O. mykiss*, age-1 individuals from genetic cluster 3, excluding Lightning Creek, were combined into a single sample (n=110). Additionally, n=15, n=34, and n=33, age-1 individuals from Little Beaver Creek, Pyramid Creek, and Stetattle Creek (genetic cluster 5) were analyzed as separate populations. The annual effective population size (N_b) from the amalgamated genetic cluster 3 was 394.9 (95 percent CI 321.0-508.5). Little Beaver Creek (genetic cluster 1) age-1 cohort had an estimated N_b of 106.2 (95 percent CI 53.8-1144.1). Pyramid Creek (genetic cluster 2) age-1 cohort had an estimated N_b of 13.5 (95 percent CI 12.5-14.7). Stetattle Creek (genetic cluster 5) age-1 cohort had an estimated N_b of 28.7 (95 percent CI 26.2-31.5).

3.2.2.4 Haplotype Diversity

The two most frequent haplotypes observed at OMY5 loci were 1-4-3-3 and 3-1-1-1 (Table 13). The haplotypes Pearse et al. (2014) found associated with juvenile propensity to emigrate (exhibit anadromous behavior) were not observed in Project *O. mykiss* (i.e., 4-3-3-1 or 1-1-1-3). The most common OMY5 haplotype overall (haplotype 3) was also the most common haplotype observed in 22 of 28 populations in the study area (data not shown).

3.2.2.5 Above and Below Gorge Dam Analysis

For population analysis of above- and below-Project (Gorge Dam) *O. mykiss*, data provided by WDFW was analyzed along with data generated by CFS. Note that *O. mykiss* data provided by WDFW has been subject to numerous previous analyses, so the intent of CFS using the data was to provide context for diversity and distance (F_{ST}) values observed, in addition to the relative magnitude of genetic differentiation. CFS requested data from upper Skagit River (below Gorge Dam), Goodell Creek, Marblemount Hatchery, lower Cascade River, and Finney Creek to incorporate into analysis representative population data for Skagit River *O. mykiss*. Importantly, the majority of WDFW data was not generated using the CRITFC-developed 354 SNPs GTSeq panel used by CFS for analysis. Rather, the WDFW data consisted of a (previous iteration) smaller 180 SNP locus panel. Therefore, CFS omitted data (nonoverlapping loci) from the total dataset for *O. mykiss* to form a complimentary set of data to use with smaller WDFW dataset. For the combined dataset, 178 loci were considered informative (minor allele frequency greater than 0.01).

Implementation of the DAPC a-score procedure on the "above-below" dataset estimated retention of 12 PCs optimized the proportion of successful reassignment corrected for the number of retained PCs (data not shown). DAPC was then rerun on the above-below *O. mykiss* dataset using 12 retained PCs, which considered 30.8 percent of the observed genetic variance in the dataset. When considering BIC based selection of various possible number of clusters (k), the primary inflection point was unclear, but $k=5$ (i.e., five genetic clusters; Figure 20) provided the highest k that both exemplified the underlying genetic principal components and resulted in reliable membership probabilities. Higher k did not change the general data pattern and merely subdivided populations on either side of above/below boundary (data not shown). Itemization of where each individual *O. mykiss* resides with respect to membership probability values are not shown here. Cluster 1 was predominantly Three Fools Creek with some Lightning Creek individuals included; most Project *O. mykiss* populations resided in cluster 2; cluster 3 was primarily Pyramid Creek; all but one individual from below Project populations resided in cluster 4, and cluster 5 was Little Beaver Creek. Rendered in 2-dimensions, the primary axis (x-axis) of Figure 20 pertains to above and below the Project. Above-Project populations (excluding Pyramid Creek) were to the left of the origin and Pyramid Creek, and below-Project populations were to the right of the origin. The second axis was driven by Little Beaver Creek genetic differentiation.

To summarize population diversity, the same location by genetic cluster designations used above were retained, with the addition of upper Skagit River, Goodell Creek, Marblemount Hatchery, lower Cascade River, and Finney Creek populations. The upper Skagit River (below Gorge Dam) and Stetattle Creek collections had the highest diversity, with below-Project populations having observed heterozygosity greater than or equal to 0.3 (data not shown). Above-Project populations had lower observed heterozygosity relative to below-Project populations. Note that the Marblemount Hatchery was the only population with greater observed heterozygosity than

expected heterozygosity, suggesting these individuals were outbred relative to Hardy-Weinberg expectations. Pairwise estimates of F_{ST} were calculated amongst populations with greater than 15 samples. The observed magnitudes of F_{ST} were distributed as expected given DAPC (Figure 21). While accounting for the highly divergent populations (Little Beaver, Three Fools, and Pyramid creeks), F_{ST} were generally larger for comparisons between above and below populations than for comparisons amongst the populations from below Gorge Dam.

3.2.2.6 Regional Analysis

A regional analysis was conducted to provide an assessment of the genetic similarity of Skagit River *O. mykiss* relative to Washington State populations from outside the Skagit River Basin. This objective was accomplished by analyzing 11,653 *O. mykiss* samples represented by 273 populations collections. This total included 30 Project and five WDFW collections from analyses described above. The collections added for this regional analysis were derived from the publicly available Columbia Basin reference genetic baseline (i.e., Columbia River Basin Mykiss GSI baseline v3.3; Hess et. al 2018). There are two caveats to this analysis. First is that while all the 180 genetic loci present in the stock identification reference genetic baseline are included in the 354 SNPs GTSeq panel used by CFS for analysis, approximately 30 percent of these loci are not present in the WDFW population data. Therefore, CFS omitted data (loci) from the total dataset for Project *O. mykiss* to form a complimentary set of data to use with WDFW data for this regional analysis. Second, the Columbia River Basin Mykiss GSI baseline v3.3 does not include data from Salish Sea populations. While these caveats could affect precision of differentiating closely related population aggregates, these data are expected to adequately resolve the primary data pattern between the coastal subspecies of *O. mykiss* (*O. m. irideus*), that is widely distributed along the Western U.S., from populations of the genetically differentiated inland subspecies of *O. mykiss* (redband; *O. m. gairdneri*).

Implementation of the DAPC a-score procedure on the regional dataset containing 11,653 individuals estimated retention of 7 PCs optimized the proportion of successful reassignment corrected for the number of retained PCs (data not shown). DAPC was then rerun on the regional *O. mykiss* dataset using 7 retained PCs, which considered 16.5 percent of the observed genetic variance in the dataset. When considering BIC based selection of various possible number of clusters (k), the primary inflection point was unclear, but k=5 (i.e., five genetic clusters; Figure 22) visualized the primary pattern underlying genetic principal components. Adopting higher k within the DAPC did not change the primary regional relationships, but subdivided populations within the coastal, interior, and study area populations (data not shown). Itemization of where each individual *O. mykiss* resides with respect to membership probability values are not shown here. Broadly speaking about populations present in DAPC analysis, 29 study area populations represented cluster 4 and one study area population (Pyramid Creek) resided in cluster 2. As seen in Figure 22, the primary axis (x-axis) was driven by variance among coastal *O. mykiss* (*O. m. irideus*) and interior redband (*O. m. gairdneri*). Project populations were placed intermediately along this axis. The secondary axis was driven by variance among Project *O. mykiss* and the 244 populations present in the reference baseline, although one of these 244 populations was Pyramid Creek. Cluster 1 was composed of one middle Columbia River and 30 lower Snake River populations. Cluster 2 was composed on Pyramid Creek (study area population), five below Project Skagit River populations, one Oregon-Washington Coastal populations, 11 lower Columbia populations, 15 Willamette River populations and 11 middle Columbia River populations (i.e.,

coastal *O. mykiss*). Cluster 3 comprised 38 lower Snake populations. Cluster 5 was composed of 47 middle Columbia River populations, eight upper Columbia populations, 13 Yakima River populations, and 63 lower Snake River populations (i.e., interior *O. mykiss*). Cluster membership for individual Skagit Basin *O. mykiss* is shown in Table 14. Project *O. mykiss* genetic characteristics appear unique compared to other populations within Washington State.

4.0 DISCUSSION

4.0 Native Char (*Salvelinus*)

The SNP analysis shows genetic structure among Bull Trout populations in the study area, with F_{ST}=0.05 indicating demographic independence (sensu Lowe and Allendorf 2010). Genetic structure is not surprising, given that Bull Trout often exhibit strong spawning site fidelity and small effective population sizes, even within the same watershed (Ardren et al. 2011; Kanda and Allendorf 1999; Spruell et al. 1999; Rieman and Allendorf 2001). The analysis showed that a significant amount of the genetic structure among Bull Trout populations could be explained by contemporary reservoir boundaries (i.e., as indicated in the AMOVA). This result suggests that reservoir-based management units (MUs) to conserve the diversity of Project area Bull Trout would have some biological basis. However, reservoir based MUs would likely provide only partial protection for the genetic diversity of Bull Trout in the Project area because unsupervised statistical models did not place Bull Trout into genetic groupings based on reservoirs. Indeed, telemetry data has provided evidence that Bull Trout can survive downstream passage potentially resulting in downstream gene flow (J. Fisher, City Light).

Bull Trout from the Project area are highly genetically distinct relative to those from downstream of Gorge Lake (i.e., lower Skagit). The genetic distinctiveness of Bull Trout from the Project area is supported by both year 1 and year 2 genetic data. A single Bull Trout sampled downstream of the Project area had a genotype that was 38 percent different compared to the 65 Bull Trout sampled from within the Project area (see PCA Figure 9). Likewise, year 1 analysis of microsatellites showed that the F_{ST} between Project area Bull Trout and populations downstream ranged from 0.207 to 0.40, which is exceptionally high, especially considering the average F_{ST} (0.32) estimated across the entire range of Bull Trout in the USA (Ardren et al., 2011). High genetic distinctiveness between Bull Trout in the Project area and those downstream could suggest the presence of unique local adaptations.

Bull Trout from the Project area are genetically less diverse relative to downstream of the Project area. The percentage of variable loci in Bull Trout from the Project area was 18 percent lower than what was observed in the single Bull Trout sampled downstream of the Project in the mainstem Skagit River in 2022. Lower diversity outside of the Project area aligns with the year 1 analysis of microsatellites, which estimated heterozygosity in downstream populations to be 37 percent higher than within the Project vicinity. Lower diversity in the Project vicinity also aligns with an unpublished mtDNA study completed by USFWS (M. Smith), which found just one mtDNA haplotype (h32) in the Project vicinity, relative to six haplotypes outside of the Project vicinity. Because the h32 haplotype has not been detected anywhere else in the Skagit basin, Smith proposed the parsimonious explanation is that the haplotype derived from the Fraser River during the Pleistocene at a time when the Skagit River was captured by the Fraser. Year 2 microsatellite

data (previously reviewed by the Expert Panel and LPs) also showed that Fraser River *O. mykiss* group closely with Fraser River fish (Blackwater, Canada) in PCA.

In a random sample of *Salvelinus* from the Project area, Bull Trout constituted 10 percent and Dolly Varden hybrids with Brook Trout were 2 percent. Brook Trout and their hybrids were more abundant in the northern part of the Project area, yet many were detected in Thunder Creek (Diablo Lake). No Brook Trout x Bull Trout hybrids were detected, but we did detect seven Brook Trout x Dolly Varden hybrids. Most (six) of the Brook Trout x Dolly Varden hybrids appeared to be first generation hybrids (F1s) with just a single fish showing evidence of back crossing towards Brook Trout. Kanda et al. (2000) observed a similar pattern, that 75 percent of hybrids between Brook Trout and Bull Trout in western Montana were F1s. By contrast, Bull Trout x Dolly Varden hybrids were substantially more common, and their distribution appeared to be multigenerational (i.e., beyond the F1 generation) and widely distributed spatially across the project. The conservation implications of hybridization between *Salvelinus* species in the Project area are nuanced because the effects of hybridization typically depend on whether it is natural or anthropogenic (i.e., "Type I" or "Type II" described in Allendorf et al. 2001). Natural hybridization between native *Salvelinus* species could be adaptive. By contrast, unnatural hybridization between invasive *Salvelinus* species is typically considered to be deleterious, resulting in wasted reproductive effort and displacement of native species and the spread of invasive traits.

Inference about the genetic diversity of Dolly Varden compared to populations across their range is unclear because reference samples of Dolly Varden from nearby populations were not available for comparison. Nevertheless, the heterozygosity at a subset of four microsatellites analyzed in collections of southern Dolly Varden from across their range was relatively similar: the range-wide HS was 0.63, compared to 0.60 in the Project vicinity (Taylor et al. 2015).

The study team observed a strong pattern of isolation-by-distance in the genetic structure of Dolly Varden in the study area. This pattern has been observed in Dolly Varden from other watersheds (Harris et al. 2015), providing a meaningful spatio-temporal model for how dispersal and gene flow tend to occur, as the distance between populations tends to be a limiting factor. However, the observed pattern suggests that over time genes can still be exchanged between more distal populations. Strong genetic structuring despite potential for exchange among populations suggests demographic independence exists among populations of Dolly Varden across the Project vicinity (sensu Lowe and Allendorf 2010). The genetic structure of Dolly Varden also suggested that there could be two major genetic populations of Dolly Varden in the Project Area, as suggested by the DAPC for k=2.

Estimates of N_e in native *Salvelinus* were small no matter how collections were grouped (e.g., by cohort or mixed) or whether estimates were corrected for bias associated with applying a model that assumes discrete generations to a species with overlapping generations. It is important to consider the quality of sampling and the scale to which the estimates of N_e apply when interpreting the results. In this study, estimates of N_e were based on linkage disequilibrium (correlation of alleles across loci) and therefore likely pertain to the local (subpopulation) spatial scale, rather than the broader metapopulation (as suggested by Whiteley et al. 2017). The localized scale of the N_e estimates is important because it suggests they may not reflect the N_e of the larger metapopulation or any other unsampled subpopulations that might contribute to the diversity of Bull Trout in the Project area, such as in Canada. Regarding sample quality, the limited sample size (65) meant that

samples had to be mixed across cohorts and aggregated into collections across sampling locations. Aggregating samples in this way violates the basic assumptions of the model. Mixing Bull Trout from different populations would be expected to produce a downward bias estimate of N_e due to the linkage disequilibrium associated with admixture (Whiteley et al. 2017).

Despite possible sampling limitations, N_e<50 suggests the potential consequences of inbreeding and loss of diversity in Bull Trout in the Project area should be taken seriously. Ardren et al. (2011) estimated N_e of Bull Trout in the Skagit river below the Project area to be greater than 200, the highest observed in the USA next to the Hoh River. Yet, they also found that 76 percent of populations across the entire North American range of Bull Trout contained effective sizes less than 50, which could mean that N_e of any given subpopulation is expected to be small regardless (e.g., in reference to the "local subpopulation scale" above). When N_e in Bull Trout is small (<50), it is recommended to conserve interconnected subpopulations that together contain at least 1,000 spawners annually (Rieman and Allendorf 2001). Importantly, the number of annual spawners is not equivalent to N_e. For example, a population with one male and 999 females contains 1,000 spawners, yet N_e is only 4. Thus, understanding the relationship between local N_e estimates and the N_e of the metapopulation in the Project area is important. The effects of subdivision on metapopulation N_e are not always obvious and can result in an N_e that is smaller or larger than the sum of the subpopulations (Allendorf et al. 2013). For example, under certain theoretical conditions, a metapopulation with isolated subpopulations can have a total N_e that is larger than their sum.

Like Bull Trout, all point estimates of N_e were less than 100 and most were less than 50 in Dolly Varden. Complicating the N_e estimates in Dolly Varden was the apparent presence of two highly divergent and widely distributed genetic populations that tended to reduce estimates of N_e within tributaries by creating LD associated with admixture. When the study team accounted for admixture by estimating N_e separately for each of the two genetic groupings within tributaries, the estimates effectively doubled. This makes sense, as combining different populations increases LD, which drives the N_e estimate lower. For instance, bias corrected N_e in Thunder Creek was 94.16 for group 1 and 73.97 for group 2 for a total N_e of 168.13 (assuming the relationship between the two groups is additive, which may not be valid). When both genetic groups were combined as a collection from a single population, N_e was 63.02.

The 50/500 rule is a general rule-of-thumb in conservation science that states N_e should not be less than 50 in the short-term, and not less than 500 in the long-term (Franklin 1980; Waples 2006; and Whiteley et al. 2012). The short-term rule is based on well-documented decreases in fitness due to inbreeding when N_e falls below approximately 50. The long-term rule is based on the loss of adaptive genetic variation that is important for potential local adaptation. In the context of a metapopulation (e.g., native trout and char in the study area), Laikre et al. (2016) recommended that long-term genetic viability should imply that the rate of inbreeding in the entire metapopulation (N_{eMeta}), as well as in the separate subpopulations (N_{eX}), should be greater than 500 due to the risk of accumulation of inbreeding within subpopulations. Complicating inferences about the threat of inbreeding due to small effective size within the study area is that hybridization with Dolly Varden (i.e., outbreeding) is common.

It is important to be cautious when making inferences about the genetic diversity, genetic structure, and effective population size of Bull Trout based on samples from only 65 individuals because the

sample size might be too small to accurately represent the entire population in the Project vicinity. Regardless, the basic inference that Bull Trout from the Project area are less diverse and are distinct compared to populations downstream is supported. Year 1 analyses provide clear evidence that the populations are not like those downstream, and the one Bull Trout analyzed from downstream of the Project area during year 2 was different, showing diversity that was multiple standard deviations above the mean, that it is unlikely to have occurred randomly. Robust Ne estimates require at least four consecutive cohort-specific estimates of Nb to obtain reliable estimates of harmonic mean Nb for a subpopulation (Whiteley et al. 2017). Nevertheless, additional samples from the Project area, from downstream, and samples from Canada would provide a more accurate representation of the genetic diversity, genetic structure, and effective population size of the Bull Trout population.

4.1 Conclusion for *Salvelinus*

The SNP analysis of Bull Trout populations in the study area showed genetic structure among them, with F_{ST}=0.05 indicating demographic independence. The analysis also showed that a significant amount of the genetic structure among Bull Trout populations could be explained by contemporary reservoir boundaries, suggesting that reservoir-based MUs to conserve diversity would have some biological basis. However, telemetry data has provided evidence that Bull Trout disperse downstream through the reservoirs, and unsupervised analysis of genetic structure demonstrated that reservoirs do not account for all the structure. Bull Trout from the Project area are highly genetically distinct from those downstream of Gorge Lake, as evidenced by the exceptionally high F_{ST} values and by the presence of haplotype h32. Bull trout in the Project area have lower genetic diversity compared to downstream populations, which is supported by SNPs, microsatellites, and mtDNA studies. Hybridization between native *Salvelinus* species was found to be more common than hybridization with invasive Brook Trout. Microsatellite analysis of Dolly Varden revealed a strong pattern of isolation-by-distance. Additionally, the genetic structure of Dolly Varden suggested that there could be two major genetic populations of Dolly Varden in the Project Area. Estimates of N_e in native *Salvelinus* were small, with most estimates less than 50. Small N_e is common in Bull Trout due to their breeding ecology, and it is recommended to conserve interconnected subpopulations at least large enough to meet 1000 spawners and/or the 50/500 rule. Metapopulations composed of multiple small populations, including subpopulations from Canada, could harbor more genetic diversity than expected due to exchange. The sample size of 65 individuals may not be sufficient to accurately represent the entire population and additional samples from the Project area, downstream, and Canada would provide a more accurate representation of the genetic diversity, genetic structure, and effective population size of the Bull Trout population.

4.2 *Oncorhynchus mykiss*

Three taxon-diagnostic loci suggested 13 percent of randomly collected *Oncorhynchus* in the Project vicinity were hybrids with Cutthroat Trout. Meaning Cutthroat and Rainbow Trout are reproducing together in the study area. Higher proportions of hybridized individuals were observed within collections from tributaries in the southern part of Ross Lake (e.g., Big Beaver).

Three study area populations were observed to be highly genetically distinct: Little Beaver Creek, Three Fools Creek, and Pyramid Creek. Passage barriers could be reenforcing genetic distinctiveness of these populations. Little Beaver Creek collections occurred above a partial

passage barrier. A partial barrier exists at the mouth of Lightning Creek, of which Three Fools is a tributary. There is a complete passage barrier present on Three Fools Creek, but collections occurred below this barrier. Meaning, individuals from Three Fools Creek had access to Lightning Creek, and some fish that were assigned to Three Fools Creek were recovered downstream in Lightning Creek collection. An analysis using the STRUCTURE program confirmed Lighting Creek individuals contained Three Fools ancestry (data not shown). Nevertheless, Three Fools Creek was distinct from adjacent Lightning Creek (and all other study area collections). Pyramid Creek collection occurred above a complete passage barrier.

In Ross Lake, Little Beaver Creek and Three Fools Creek populations are distinct and probably demographically independent, given their high degree of genetic differentiation (Sensu Lowe and Allendorf 2010).

Reproductive connection with Three Fools Creek is likely the source of Lightning Creek distinctiveness (STRUCTURE analysis, data not shown).

There is a complete passage barrier at the mouth of Pyramid Creek. Pyramid Creek fish were genetically aberrant in that this population appears to be derived from the coastal *O. mykiss* lineage (*O. m. irideus* subspecies). The coastal lineage is distributed widely along the Western U.S. and is the lineage of lower Skagit River *O. mykiss* (below Gorge Dam).

The two genetic clusters widely distributed in the study area (labelled cluster 3 and cluster 5) were stable across multiple years of collections (collections were combined across years for a location). As reproductive connection can homogenize genetic diversity within a short time (i.e., years), these clusters must be persisting through (non-random) assortative mating with respect to cluster identity.

The two genetic clusters distributed throughout the study area (labelled cluster 3 and cluster 5) were differentiated from each other, even for collections from the same location. For example, Granite Creek-3 and Canyon Creek-3 had a smaller distance (F_{ST} =0.005) between them than either Granite Creek-3 to Granite Creek-5 (distance=0.030) or Canyon-3 to Canyon-5 (distance=0.035). This pattern held for comparisons among lakes, with smaller distances (F_{ST}) observed between cluster-3 populations across Ross, Diablo, and Gorge Lakes (Ruby, Canyon, Granite, Colonial, Thunder, Stetattle Creeks) than comparisons between cluster 3 and 5 from within the same tributary.

Study results indicate that the current classification that there is a single population in the study area is not accurate.

Categorizing the tributaries based on genetic distance (e.g., F_{ST}~0.02) is challenging given the presence of two genetic clusters (labeled 3 and 5). Considering just genetic cluster 3, the study team suggests that these collections exhibit enough similarity to be treated as a single management unit, despite not having completely random mating. Yet, it seems unwise to ignore clusters 3 and 5, given the unknown qualities of their differences and dynamics of persistence.

Given study area non-migratory (resident) *O. mykiss* are not protected under the state or federal Endangered Species Acts, and management decisions have considered the Project a single population, the study team proposes initially classifying as four populations in Project: 1) Little

Beaver Creek, 2) Three Fools Creek, 3) Lightning Creek, and 4) the remaining tributaries (excluding Pyramid Creek). At some future timepoint, management may have to account for reproductive dynamics between clusters 3 and 5 (at same location), which could necessitate altering the classification of populations to location by genetic cluster.

Haplotype diversity was observed at chromosome 5 loci (OMY5), a location potentially associated with juvenile life-history. Future evaluations of adaptive diversity could be conducted with data generated from this study and compared to other geographic regions. These study data could inform future deliberations regarding quantitative trait diversity present in the study area.

Genetic distances were considerably higher for comparisons between study area *O. mykiss* with collections representing populations from below Gorge Dam (upper Skagit River, Goodell Creek, lower Cascade River, Finney Creek, Marblemount Hatchery). Genetic distances (F_{ST}) were approximately an order of magnitude larger for tests between above Gorge Dam to populations below Gorge Dam. It is reasonable to conclude that Project *O. mykiss* are distinct from Skagit River *O. mykiss*. Pairwise F_{ST} estimates among below-Project populations were less than 0.02, although the Marblemount Hatchery collection was more divergent, with F_{ST} = 3.5-5.1 when compared to other populations from below Gorge Dam.

Project *O. mykiss* had lower genetic diversity than that observed for below-Project populations, although not remarkably so. The magnitude of genetic diversity present in Project *O. mykiss* does not appear to be a concern, especially considering the reproductive connectivity observed within Project *O. mykiss* populations.

A regional comparison was made for Project *O. mykiss* using a dataset consisting of 243 collections of *O. mykiss*. In total, 11,653 *O. mykiss* samples were included in this analysis. The results observed were similar to observations reported for the 2019 stranded *O. mykiss* analysis (Small et al. 2020). Project *O. mykiss* were intermediate (on first genetic principal component) with respect to coastal (*O. m. irideus*) and inland redband (*O. m. gairdneri*) ancestry. Additionally, the second genetic principal component represented genetic variation between Project *O. mykiss* and all other populations present in reference database. Project *O. mykiss* genetic characteristics appear unique compared to other populations within Washington State.

The uniqueness of Project *O. mykiss* has implications for discussions regarding human-mediated passage.

4.3 Conclusions

Study area *O. mykiss* had lower diversity than that observed for populations from below Gorge Dam, but not remarkably so. The annual effective number of breeders (N_b) was estimated from some study area locations where samples numbers were sufficient for 2022 age-1 cohort. Initial estimates of N_b were calculated from the widely distributed genetic cluster 3 (excluding Lightning Creek), Little Beaver Creek, Pyramid Creek, and Stetattle Creek genetic cluster 5. The N_b from the amalgamated genetic cluster 3 was 394.9 (95 percent CI 321.0-508.5). Little Beaver Creek (genetic cluster 1) age-1 cohort had an estimated N_b of 106.2 (95 percent CI 53.8-1144.1). Pyramid Creek (genetic cluster 2) age-1 cohort had an estimated N_b of 13.5 (95 percent CI 12.5-14.7). Stetattle Creek (genetic cluster 5) age-1 cohort had an estimated N_b of 28.7 (95 percent CI 26.2-31.5). Given reproductive connectivity, lower genetic diversity observed is not a concern.

Three study area populations were observed to be highly genetically distinct: Little Beaver Creek, Three Fools Creek, and Pyramid Creek. Two additional genetic clusters widely distributed throughout the study area were differentiated from each other, even for collections obtained from the same location. The study results indicate that the current classification that there is a single population in the study area is not accurate. The study team proposes initially classifying *O. mykiss* in the study area as: (1) Little Beaver Creek; (2) Three Fools Creek; (3) Lightning Creek; and (4) the remaining tributaries (excluding Pyramid Creek).

Genetic distances were an order of magnitude higher for comparisons between study area *O. mykiss* with collections representing populations from below Gorge Dam. It is reasonable to conclude that Project *O. mykiss* are distinct from Skagit River *O. mykiss*. Pyramid Creek fish were genetically aberrant in that this population appeared to be derived from the coastal *O. mykiss* lineage (i.e., from below Gorge Dam). Study area *O. mykiss* appear unique compared to other populations from Washington State.

5.0 REFERENCES

Allendorf, F.W., and S.R. Phelps. 1981. Use of Allelic Frequencies to Describe Population Structure. Canadian Journal of Fisheries and Aquatic Sciences 38(12):1507–1514.

Allendorf, F.W., R.F. Leary, P. Spruell, and J.K. Wenburg. 2001. The problems with hybrids: setting conservation guidelines. Trends in ecology & evolution, 16(11), pp.613-622.

Allendorf, F.W., G. Luikart, and S.N. Aitken. 2013. Conservation and the genetics of populations. Wiley-Blackwell.

Ardren, W.R., P.W. DeHaan, C.T. Smith, E.B. Taylor, R. Leary, C.C. Kozfkay, D.K. Hawkins. (2011). Genetic Structure, Evolutionary History, and Conservation Units of Bull Trout in the Coterminous United States. Transactions of the American Fisheries Society, 140(2), 506–525.

Bohling, J., J. Von Bargen, M. Piteo, A. Louden, M. Small, T.A. Delomas, and R. Kovach. 2021. Developing a Standardized Single Nucleotide Polymorphism Panel for Rangewide Genetic Monitoring of Bull Trout. North American Journal of Fisheries Management 41(6):1920–1931.

Campbell, N.R., S.A. Harmon, and S.R. Narum. 2014. Genotyping-in-Thousands by sequencing (GTseq): A cost effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources 15(4):855–867.

Copeland, T., K. Hernandez, M.T. Davison, and K. Wright. 2018. Validation of Spawn Checks and Saltwater Age Assignments Based on Scales from Known Repeat-Spawning Steelhead. North American Journal of Fisheries Management 38(5):1050–1058.

Fisher, J. 2022. Seattle City Light correspondence.

Franklin, I.R. 1980. Evolutionary change in small populations. Pages 135–150 Conservation Biology An Evolutionary-Ecological Perspectiveeds Soule ME, Wilcox BA. Sinauer, Sunderland, Massachusetts.

Garant, D., J.J. Dodson, and L. Bernatchez. 2000. Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar L.*). Molecular Ecology 9(5):615–628.

Harris, L.N., R. Bajno, C.P. Gallagher, I. Koizumi, L.K. Johnson, K.L. Howland, E.B. Taylor, and J.D. Reist. 2015. Life-history characteristics and landscape attributes as drivers of genetic variation, gene flow, and fine-scale population structure in northern Dolly Varden (*Salvelinus malma malma*) in Canada. Canadian Journal of Fisheries and Aquatic Sciences.

Hernandez, K., T. Copeland, and K. Wright. 2014. Quantitative Assessment of Scale Resorption in Migrating and Spawning Steelhead of the Snake River Basin. Transactions of the American Fisheries Society 143(6):1562–1568.

Hess, J.E., S.A. Harmon, D.J. Hasselman, R.L. Horn, A.R. LaGrange, A.P. Matala, and S.R. Narum. 2018. Genetic Assessment Of Columbia River Stocks, 1/1/2018 - 12/31/2018. Annual Report for Bonneville Power Administration, 2008-907-00.

Hutchison, D.W., and A.R. Templeton. 1999. Correlation of Pairwise Genetic and Geographic Distance Measures: Inferring the Relative Influences of Gene Flow and Drift on the Distribution of Genetic Variability. Evolution 53(6):1898–1914.

Isermann, D.A., and C.T. Knight. 2005. A Computer Program for Age–Length Keys Incorporating Age Assignment to Individual Fish. North American Journal of Fisheries Management 25(3):1153–1160.

Jombart, T. 2011. *adegenet*: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403-1405.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27(21):3070–3071.

Kanda, N. and Allendorf, F.W., 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(1), pp.92-106.

Kanda, N., R.F. Leary, and F.W. Allendorf. 2002. Evidence of introgressive hybridization between bull trout and brook trout. Transactions of the American Fisheries society, 131(4), pp.772-782.

Laikre, L., F. Olsson, E. Jansson, O. Hössjer, and N. Ryman. 2016. Metapopulation effective size and conservation genetic goals for the Fennoscandian wolf (*Canis lupus*) population. Heredity 117(4):279–289.

Love, D.C. 2016. Manual for aging steelhead trout scales based on scale sampling from Sitkoh Creek, Big Ratz Creek, and other Southeast Alaska streams, 2003-2011. Alaska Department of Fish and Game, Division of Sport Fish, Research and Technical Services.

Lowe, W.H., and F.W. Allendorf. 2010. What can genetics tell us about population connectivity? Molecular Ecology 19(15):3038–3051.

Melnik, N. O., G. N. Markevich, E. B. Taylor, A. V. Loktyushkin, and E. V. Esin. 2020. Evidence for divergence between sympatric stone charr and Dolly Varden along unique environmental gradients in Kamchatka. Journal of Zoological Systematics and Evolutionary Research 58(4):1135–1150.

Murphy, B.R., and D.W.E. Willis. 1996. Fisheries Techniques 2nd Edition. American Fisheries Society, Bathesda, MD.

Ogle, D.H., J.C. Doll, P. Wheeler, and A. Dinno. 2022. FSA: Fisheries Stock Analysis. R package version 375 0.9. 3.

Ostberg, C.O., and D.M. Chase. 2022. Ontogeny of eDNA shedding during early development in Chinook Salmon (*Oncorhynchus tshawytscha*). Environmental DNA, 4(2), 339-348.

Pearse, D.E., M.R. Miller, A. Abadía-Cardoso, and J.C. Garza. 2014. Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. Proceedings of the Royal Society B: Biological Sciences 281(1783):20140012.

Reynolds, J.B., A.L. Kolz, A.V. Zale, D.L. Parrish, and T.M. Sutton. 2012. Electrofishing. Pages 305–361.

Rice, W.R., 1989. Analyzing tables of statistical tests. Evolution, 43(1), pp.223-225.

Rieman, B.E. and F.W. Allendorf. 2001. Effective population size and genetic conservation criteria for bull trout. North American Journal of Fisheries Management 21:756-764.

Ryman, N., L. Laikre, and O. Hössjer. 2019. Do estimates of contemporary effective population size tell us what we want to know? Molecular Ecology 28(8):1904–1918.

Seamons, T. 2020. Washington Department of Fish and Wildlife (WDFW) correspondence.

Seattle City Light (City Light). 2021. Revised Study Plan (RSP) for the Skagit River Hydroelectric Project, FERC Project No. 553. April 2021.

Small, M.P., E. Lowery, and M. Kissler. 2020. Genetic analysis of Char from Gorge Lake, Upper Skagit River, Washington. Washington Dept. of Fish and Wildlife Final Report.

Sinnwell, J.P., D.J. Schaid, and M.J.P. Sinnwell. 2022. Package 'haplo. stats.'

Spruell, P., B.E. Rieman, K.L. Knudsen, F.M. Utter, and F.W Allendorf. 1999. Genetic population structure within streams: microsatellite analysis of bull trout populations. Ecology of Freshwater Fish, 8(3), 114–121. doi:10.1111/j.1600-0633.1999.tb00063.x.

Taylor, E.B., and S.L. May-McNally. 2015. Genetic analysis of Dolly Varden (*Salvelinus malma*) across its North American range: evidence for a contact zone in southcentral Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 72(7), 1048–1057.

Taylor, E.B., M.D. Stamford, and J.S. Baxter. 2003. Population subdivision in westslope cutthroat trout (*Oncorhynchus clarki lewisi*) at the northern periphery of its range: evolutionary inferences and conservation implications. Molecular Ecology, 12(10), 2609–2622.

Waples, R.S. 2006. A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics (7):167–184.

Waples, R.S., and C. Do. 2008. ldne: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8(4):753–756.

Waples, R.S., T. Antao, and G. Luikart. 2014. Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. Genetics 197(2):769–780.

Whiteley, A.R., J.A. Coombs, M. Hudy, Z. Robinson, K.H. Nislow, and B.H. Letcher. 2012. Sampling strategies for estimating brook trout effective population size. Conservation Genetics 13(3):625–637.

Whiteley, A.R., J.A. Coombs, M.J. O'Donnell, K.H. Nislow, and B.H. Letcher. 2017. Keeping things local: Subpopulation Nb and Ne in a stream network with partial barriers to fish migration. Evolutionary Applications, 10(4), pp.348-365.

Wright, S., 1949. The genetical structure of populations. Annals of eugenics, 15(1), pp.323-354.

6.0 TABLES

Table 1. Summary of locations[1] targeted for sampling of *Oncorhynchus* and *Salvelinus* during the 2022 field season.

No.	Skagit River Drainage	River/Stream Name	Location Description	Sampled in 2022
1	Ross Lake	Hozomeen Creek	Mainstem	Yes
2		Freezeout Creek	Mainstem	No
3		Lightning Creek	Mainstem	Yes
4		Three Fools Creek	Mainstem	Yes
5		Cinnamon Creek	Mainstem	No
6		Castle Fork	Mainstem	No
7		Devils Creek	Mainstem and tributaries	No
8		North Fork Devils Creek	Mainstem	No
9		Roland Creek	Mainstem	Yes
10		Ruby Creek	Mainstem	Yes
11		Canyon Creek	Mainstem up to cascade barrier	Yes
12		North Fork Canyon Creek	Mainstem	No
13		Granite Creek	Mainstem up to cascade barrier	Yes
14		Panther Creek	Mainstem up to cascade barrier	Yes
15		Pierce Creek	Mainstem	Yes
16		Big Beaver Creek	Mainstem and tributaries including Beaver Ponds	Yes
17		McMillan Creek	Mainstem	Yes
18		Luna Creek	Mainstem	Yes
19		Little Beaver Creek	Mainstem above and below barriers	Yes
20		Silver Creek	Mainstem	Yes
21	Diablo Lake	Thunder Creek	Mainstem	Yes
22		Colonial Creek	Mainstem	Yes
23	Gorge Lake	Stetattle Creek	Mainstem above and below barrier	Yes
24		Pyramid Creek	Mainstem	Yes
25		Gorge Creek	Mainstem	No

1 The proposed sampling plan for 2022 field study, including sampling locations, was reviewed and approved by the Expert Panel and interested LPs. Locations include those sampled by CFS and USGS.

Table 2. Diagnostic SNP markers for taxa within *Salvelinus*. These markers were used to distinguish among Bull Trout, Brook Trout, and Dolly Varden.

Assay Name	Bull Trout	Brook Trout	Dolly Varden
Salv_SNP_008	A:A	T:T	A:A
Salv_SNP_013	T:T	G:G	G:G
sf000151_AT_HYB	T:T	C:C	C:C
sf000157_01AT_HYB	A:A	C:C	C:C
sf000382_AG_HYB	T:T	A:A	T:T
sf000508_CT_HYB	C:C	G:G	0
sf000559_AG_HYB	A:A	G:G	G:G
sf000754_AC_HYB	C:C	G:G	C:C
sf001164_02GT_HYB	T:T	A:A	A:A
sf002131_AG_HYB	A:A	G:G	A:A
sf002792_01AG_HYB	G:G	C:C	G:G
sf003611_AC_HYB	A:A	G:G	A:A
sf004651_AG_HYB	G:G	A:A	0
sf005440_AG_HYB	G:G	C:C	0
Sfo_12199_79192_HYB	T:T	A:A	A:A
Sfo_2714_25693_HYB	A:A	T:T	A:A*
Sfo_3881_34908_HYB	T:T	A:A	T:T
Sfo_4699_39079_HYB	T:T	A:A	A:A
Sfo_4701_39083_HYB	A:A	C:C	C:C
Sfo_5504_43035_HYB	C:C	T:T	C:C
Sfo_579_12874_HYB	T:T	C:C	T:T
Sfo_9883_66689_HYB	T:T	C:C	T:T

1 * Only amplified in Dolly Varden about 50 percent of the time.

Table 3. Summary statistics[1] of 2022 Bull Trout collections.

Collection pool[1]	N	H_o	H_s	HWP	F_{IS}	N_e
Study Area	65	0.29(0.16)	0.23(0.17)	1/33	-0.26	31.40 (17.50, 68.80)
Ross Lake	39	0.28(0.16)	0.29(0.16)	0/33	0.03	30.9 (17.3, 90.6)
Diablo Lake	3	NA	NA	NA	NA	NA
Gorge Lake	23	0.34(0.17)	0.33(0.15)	1/28	-0.03	10.9 (6.1, 22.4)
inferred1	31	0.32(0.18)	0.31(0.15)	0/30	-0.03	98.50 (26.20, infinite)
inferred2	13	0.43(0.20)	0.38(0.15)	0/24	-0.13	6.7 (2.60, 27.0)
inferred3	21	0.28(0.17)	0.27(0.15)	0/32	-0.04	24.50 (12, 173)

1 H_S = expected heterozygosity, H_O = observed heterozygosity, HWP = number of markers in significant deviation from Hardy-Weinberg Proportions (α=0.05), F_{IS} = multilocus deviation from expected heterozygosity, N_e = Effective population size (unadjusted mixed cohort). 'NA' indicates sample size was too small to estimate the parameter. Numbers in parentheses are standard deviation except for in the Ne column, in which case parentheses contain the parametric 95 percent CI.

2 Each row represents a distinct collection pool because samples were too small within individual tributaries to be treated separately.

Table 4. **Summary statistics for collections of Dolly Varden (N=413) and for the 12 inferred genetic groupings identified by DAPC (N=405). Only individuals with greater than 0.50 probability of assignment to an inferred cluster were included (N=405).**

Collection	N	H_o	H_s	HWP	F_{IS}	N_e
Big Beaver	43	0.67 (0.27)	0.67 (0.29)	1/8	-0.01	30.5 (23.4, 41.3)
Canyon	47	0.63 (0.27)	0.64 (0.29)	2/8	0.00	25.1 (19.9, 32.1)
Colonial	22	0.74 (0.17)	0.72 (0.21)	2/8	-0.05	20.7 (13.8, 34.6)
Granite	22	0.64 (0.35)	0.64 (0.35)	0/8	-0.01	24.1 (15.5, 43.5)
Hozomeen	3	NA	NA	NA	NA	NA
Lightning	136	0.55 (0.26)	0.58 (0.28)	2/8	0.04	21.3 (18.1, 24.9)
NF Canyon	4	NA	NA	NA	NA	NA
Pierce	1	NA	NA	NA	NA	NA
Roland	1	NA	NA	NA	NA	NA
Ruby	29	0.59 (0.32)	0.63 (0.35)	2/8	0.04	21.3 (15.3, 31.6)
Silver	6	NA	NA	NA	NA	NA
Stetattle	7	NA	NA	NA	NA	NA
Thunder	92	0.68 (0.28)	0.69 (0.26)	0/8	0.02	34.4 (29.2, 40.7)
Inferred 1	39	0.42(0.26)	0.62(0.23)	3/8	0.27	24.2 (16.9, 36.8)
Inferred 2	43	0.62(0.26)	0.66(0.28)	2/8	0.05	26.0 (20.5, 33.6)
Inferred 3	27	0.67(0.30)	0.65(0.29)	0/8	-0.04	26.2 (18.3, 41.2)
Inferred 4	27	0.67(0.29)	0.65(0.27)	0/8	-0.03	17.8 (12.0, 28.2)
Inferred 5	37	0.62(0.26)	0.67(0.28)	3/8	0.06	33.5 (25.6, 46.0)
Inferred 6	38	0.64(0.32)	0.64(0.32)	0/8	-0.01	28.7 (22.0, 38.7)
Inferred 7	33	0.67(0.34)	0.66(0.31)	0/8	0.01	36.2 (25.9, 54.7)
Inferred 8	41	0.61(0.27)	0.64(0.29)	1/8	0.05	31.8 (23.7, 44.8)
Inferred 9	36	0.68(0.26)	0.67(0.28)	1/8	-0.02	28.3 (21.5, 38.6)
Inferred 10	43	0.62(0.32)	0.64(0.32)	1/8	0.03	34.2 (26.0, 47.0)
Inferred 11	15	0.64)0.32)	0.66(0.30)	0/8	0.03	20.1 (11.4, 48.5)
Inferred 12	26	0.71(0.23)	0.69(0.23)	1/8	-0.05	31.0 (20.9, 52.3)

1 Notes: H_S = expected heterozygosity; H_O = observed heterozygosity; HWP = number of markers in significant deviation from Hardy-Weinberg Proportions (α=0.05); F_{IS} = multilocus deviation from expected heterozygosity; N_e = Effective population size (unadjusted mixed cohort).

2 'NA' indicates sample size was too small to estimate the parameter. Numbers in parentheses are standard deviation except for the N_e column where parentheses contain the parametric 95 percent CI.

Table 5. Non-hybrid *O. mykiss* samples used for genetic analysis.

No.	River/Stream Name	Collection Years[2]	Sample Size (*n*)
1	Silver Creek	2022	44
2	Hozomeen Creek	2022	3
3	Little Beaver Creek	2019, 2021, 2022	73
4	Lightning Creek	2018, 2020, 2022	143
5	Three Fools Creek	2019, 2020, 2021, 2022	103
6	Big Beaver Creek	2019, 2022	52
7	McMillan Creek	2021, 2022	6
8	Pierce Creek	2022	3
9	Roland Creek	2022	8
10	Ross Lake	2020, 2022	4
11	Ruby Creek	2019, 2022	185
12	Canyon Creek	2018, 2019, 2022	94
13	North Fork Canyon Creek	2022	6
14	Panther Creek	2022	35
15	Granite Creek	2019, 2022	109
16	Colonial Creek	2022	53
17	Thunder Creek	2019, 2022	98
18	Stetattle Creek	2021, 2022	125
19	Pyramid Creek	2022	60
20	Gorge Lake[1]	2019	30

1 WDFW genotype data.
2 CFS collections from 2022. All other dates shown are USGS collected samples.

Table 6. Non-hybrid *O. mykiss* samples used for genetic analysis. Bold F_{IS} values were statistically significant from zero.

No.	Location/DAPC cluster[1]	n	F_{IS}	Allelic Richness
1	Silver Creek-3	44	**0.058**	1.303
2	Hozomeen Creek-3	3	**-0.067**	NA
3	Little Beaver Creek-1	73	**0.062**	1.161
4	Lightning Creek-3	138	**0.04**	1.269
5	Three Fools Creek-4	108	0.036	1.116
6	Big Beaver Creek-3	46	0.021	1.291
7	Big Beaver Creek-5	6	-0.057	1.271
8	McMillan Creek-3	6	**0.104**	NA
9	Pierce Creek-3	3	**0.061**	NA
10	Roland Creek-3	8	0.008	NA
11	Ross Lake-3	4	-0.111	1.309
12	Ruby Creek-3	174	0.019	1.299
13	Ruby Creek-5	11	0.026	1.278
14	Canyon Creek-3	47	**0.032**	1.294
15	Canyon Creek-5	47	**0.034**	1.276
16	NF Canyon Creek-5	6	-0.023	NA
17	Panther Creek-3	3	-0.071	1.268
18	Panther Creek-5	32	0.016	1.222
19	Granite Creek-3	41	**0.022**	1.287
20	Granite Creek-5	68	**0.052**	1.266
21	Colonial Creek-3	53	0.015	1.299
22	Thunder Creek-3	96	**0.023**	1.303
23	Thunder Creek-5	2	**-0.14**	NA
24	Stetattle Creek-3	59	0.022	1.306
25	Stetattle Creek-5	66	**0.049**	1.321
26	Pyramid Creek-2	60	**0.048**	1.252
27	Gorge Lake-3 [2]	29	0.01	NA
28	Gorge Lake-5 [2]	1	NA	NA

1 Populations are label by location description and DAPC cluster membership.
2 WDFW genotype data.

Table 7. **Pairwise F_{ST} for Bull Trout collection pools based on reservoirs and inferred genetic clusters from DAPC analysis. Significance at the alpha=0.05 level is indicated by bold lettering.**

	Diablo Lake	Gorge Lake	Ross Lake	inferred1	inferred2	inferred3
Diablo Lake	0					
Gorge Lake	0.04729	0				
Ross Lake	0.02634	**0.02623**	0			
inferred1	0.0052	**0.02554**	**0.0226**	0		
inferred2	**0.13765**	**0.04016**	**0.09551**	**0.14063**	0	
inferred3	**0.09041**	**0.08098**	0.01246	**0.09945**	**0.15103**	0

Table 8. **Size and age summaries for aged Salvelinus individuals.**

Age Class	n	Min	Max	Mean	SD
0+	33	60	118	78.97	12.95
1+	60	76	188	123.41	21.4
2+	9	160	256	219.18	24.2
3+	7	323	426	356.25	35.31
5+	1	420	420	420	0

Table 9. **Size and age summaries for aged and age-assigned *Salvelinus* individuals.**

Age Class	*n*	Min	Max	Mean	SD
0+	98	60	118	78.97	12.95
1+	175	76	188	123.41	21.4
2+	28	160	256	219.18	24.2
3+	8	323	426	356.25	35.31
5+	2	420	420	420	0

Table 10. **Effective population size estimates for Bull Trout corrected for overlapping generations using Waples et al. (2014) adjustment based on adult life span (8.5 years) and age of first reproduction (3.0) (Hemmingsen et al. 2001). Note: Bias corrections are insensitive to Adult Lifespan and Age of Maturity within a few years. Only collections with greater than 20 samples were analyzed.**

Population	Cohort	Age Class	N_b Raw	Lower	Upper	Adult Lifespan	Age at Maturity	N_b Adjusted	N_e Adjusted
Study Area	2021-	0 to 1+	6.4	3.2	10.4	8.5	3	5.69	12.08
Study Area	2020-	2+	47.2	23.4	194	8.5	3	42.68	88.35

Table 11. Effective population size estimates for Dolly Varden corrected for overlapping generations using Waples et al. (2014) adjustment based on adult life span and age of first reproduction.

Population	Cohort	Age Class	N_b Raw	Lower	Upper	Adult Lifespan	Age at Maturity	N_b Adjusted	N_e Adjusted
Big Beaver	2020	1	23.00	16.90	32.7	8.5	2.5	20.72	43.13
Lightning	2021	0	14.50	10.00	22.1	8.5	2.5	13.02	27.24
Lightning	2020	1	24.70	18.00	35.4	8.5	2.5	22.26	46.31
Thunder	2021	0	33.70	26.6	44.10	8.5	2.5	30.42	63.13
Thunder	2020	1	27.90	17.3	56.6	8.5	2.5	25.16	52.29
Lightning_2*	2021	0	28.4	16.6	64.8	8.5	2.5	25.62	53.22
Lightning_2*	2020	1	24.6	16.3	40.5	8.5	2.5	22.17	46.12
Thunder_1*	2021	0	50.3	25.5	259	8.5	2.5	45.47	94.16
Thunder_2*	2021	0	39.5	23.5	92.7	8.5	2.5	35.68	73.97

Adult Lifespan and Age of Maturity taken from Jonsson et al. (1984). Only collections with greater than 20 samples were analyzed.

*indicates *a posteriori* grouping of fish into collections that may represent two highly divergent populations.

Table 12. Size and age summaries for aged and age-assigned Oncorhynchus individuals.

age	n	mean	sd	min	Q1	median	Q3	max
0+	86	80.05	14.35	55	68	80.5	91.75	104
1+	248	117.76	23.27	69	101	116	133.25	174
2+	126	162.98	26.08	106	144	163	182.75	213
3+	34	257.5	36.07	171	225	260	289.5	320
4+	12	339.17	26.79	305	312.25	347.5	360	380
5+	2	341	1.41	340	340.5	341	341.5	342

Table 13. Diversity of OMY5 haplotypes for Project *O. mykiss* individuals.

Haplotype	OmyR14589	OmyR19198	OmyR24370	OmyR33562	Hap.freq
1	1	1	1	1	0.000
2	1	4	1	3	0.002
3	1	4	3	3	0.867
4	3	1	1	1	0.123
5	3	1	1	3	0.001
6	3	4	1	3	0.007
7	3	4	3	3	0.000

Table 14. Counts of DAPC cluster membership from regional analysis showing only Skagit Basin *O. mykiss* samples used for regional genetic analysis. Population labels were retained from Table 6 and clusters are identical to Figure 22. Clusters-1, -3, and -5 pertain to redband subspecies, cluster-2 pertains to coastal subspecies, and cluster-4 pertains to study area *O. mykiss*.

Population	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Silver Creek-1	0	0	0	1	0
Silver Creek-3	0	0	0	43	1
Hozomeen Creek-3	0	0	0	2	1
Little Beaver Creek-5	0	0	0	73	0
Lightning Creek-1	0	0	0	2	0
Lightning Creek-3	0	0	0	137	1
Three Fools Creek-2	0	0	0	106	0
Big Beaver Creek-1	0	0	0	3	0
Big Beaver Creek-3	0	0	0	49	0
McMillan Creek-3	0	0	0	6	0
Pierce Creek-1	0	0	0	1	0
Pierce Creek-3	0	0	0	2	0
Roland Creek-1	0	0	0	2	0
Ross Lake-3	0	0	0	9	1
Ruby Creek-1	0	0	0	14	0
Ruby Creek-3	0	0	0	170	1
Canyon Creek-1	0	0	0	35	1
Canyon Creek-3	0	0	0	58	0
NF Canyon Creek-1	0	0	0	6	0
Panther Creek-1	0	0	0	17	0
Panther Creek-3	0	0	0	18	0
Granite Creek-3	0	1	0	108	0
Colonial Creek-1	0	0	0	4	0
Colonial Creek-3	0	0	0	49	0
Thunder Creek-1	0	0	0	12	0
Thunder Creek-3	0	0	0	86	0
Stetattle Creek-1	0	3	0	64	2
Stetattle Creek-3	0	0	0	56	0
Pyramid Creek-4	0	59	0	1	0
Gorge Lake-3 [1]	0	0	0	30	0
upper Skagit [1]	0	147	0	1	0
Finney Creek [1]	0	53	0	0	0
Goodell Creek [1]	0	99	0	0	0
lower Cascade [1]	0	20	0	0	1
Marblemount [1]	0	106	0	0	0

1 WDFW genotype data.

7.0 FIGURES

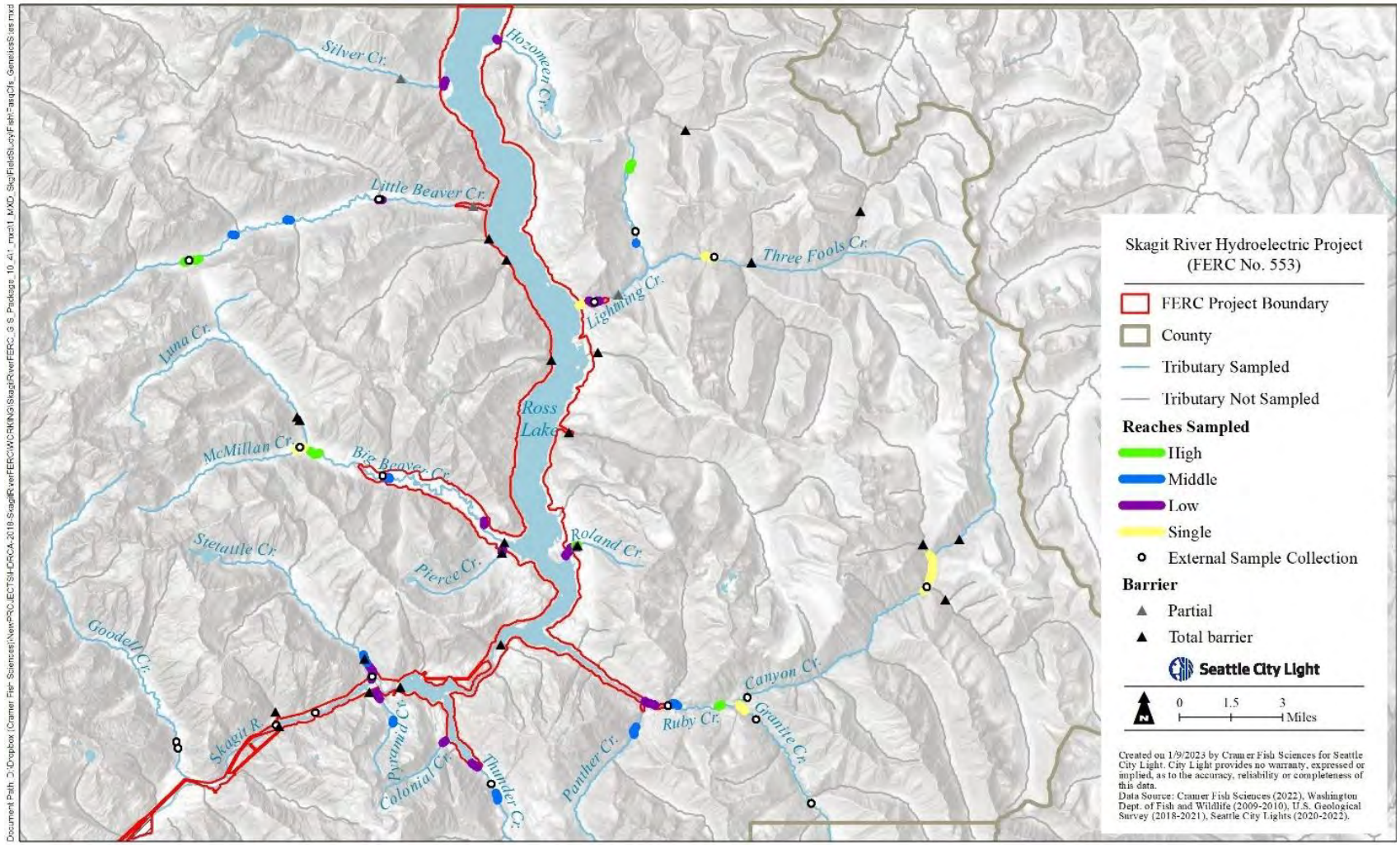


Figure 1. **Year 2 *Salvelinus* and *Oncorhynchus* genetics sampling locations. The study team attempted to sample high, mid, and low elevation reach on all primary tributaries and a single low reach on all secondary tributaries. Sample locations chosen based on specific LP and EP request.**

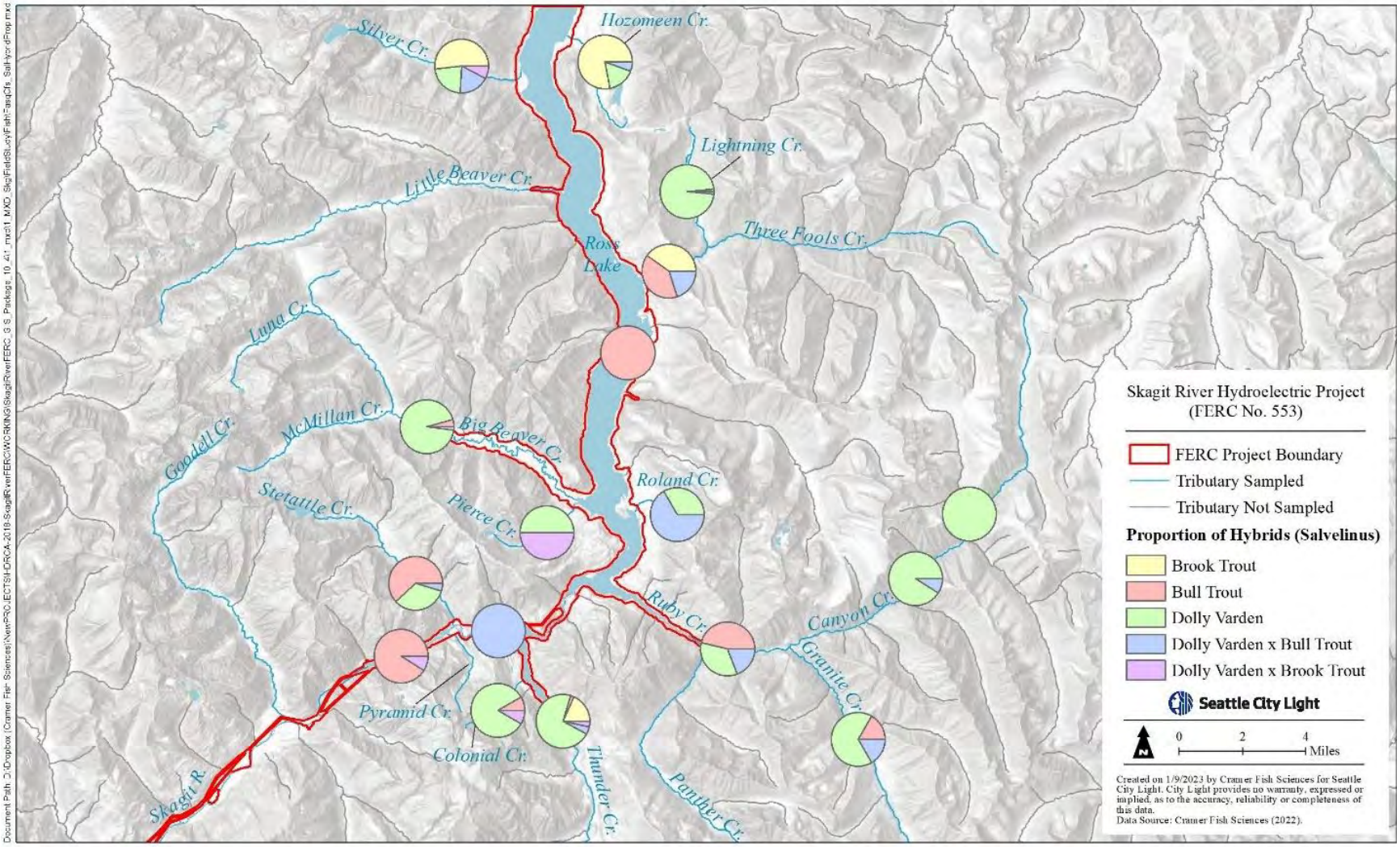


Figure 2. **Map of Year 2 *Salvelinus* collections showing the proportion of individuals that were Bull Trout, Dolly Varden, Brook Trout, or hybrids based on 20 taxon-diagnostic SNPs.**

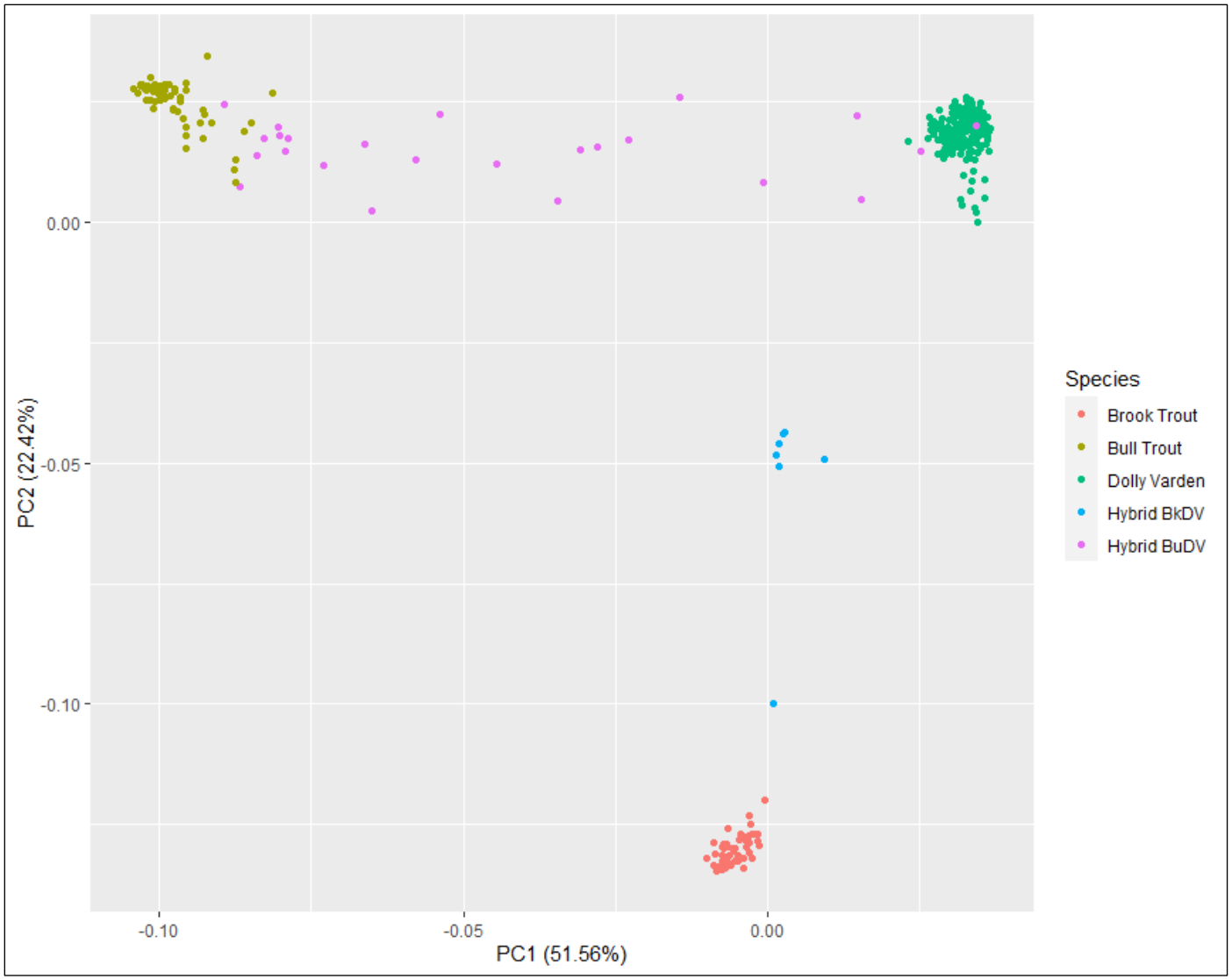


Figure 3. **Scatterplot of first 2 PCs based on genotypes at 263 GTseq SNPs within Salvelinus collected in the study area, 2022. The plot shows clear distinction among Brook Trout, Bull Trout, and Dolly Varden, and also highlights the presence of hybrids. Individuals colored as hybrids were identified using taxon-diagnostic markers not PC scores.**

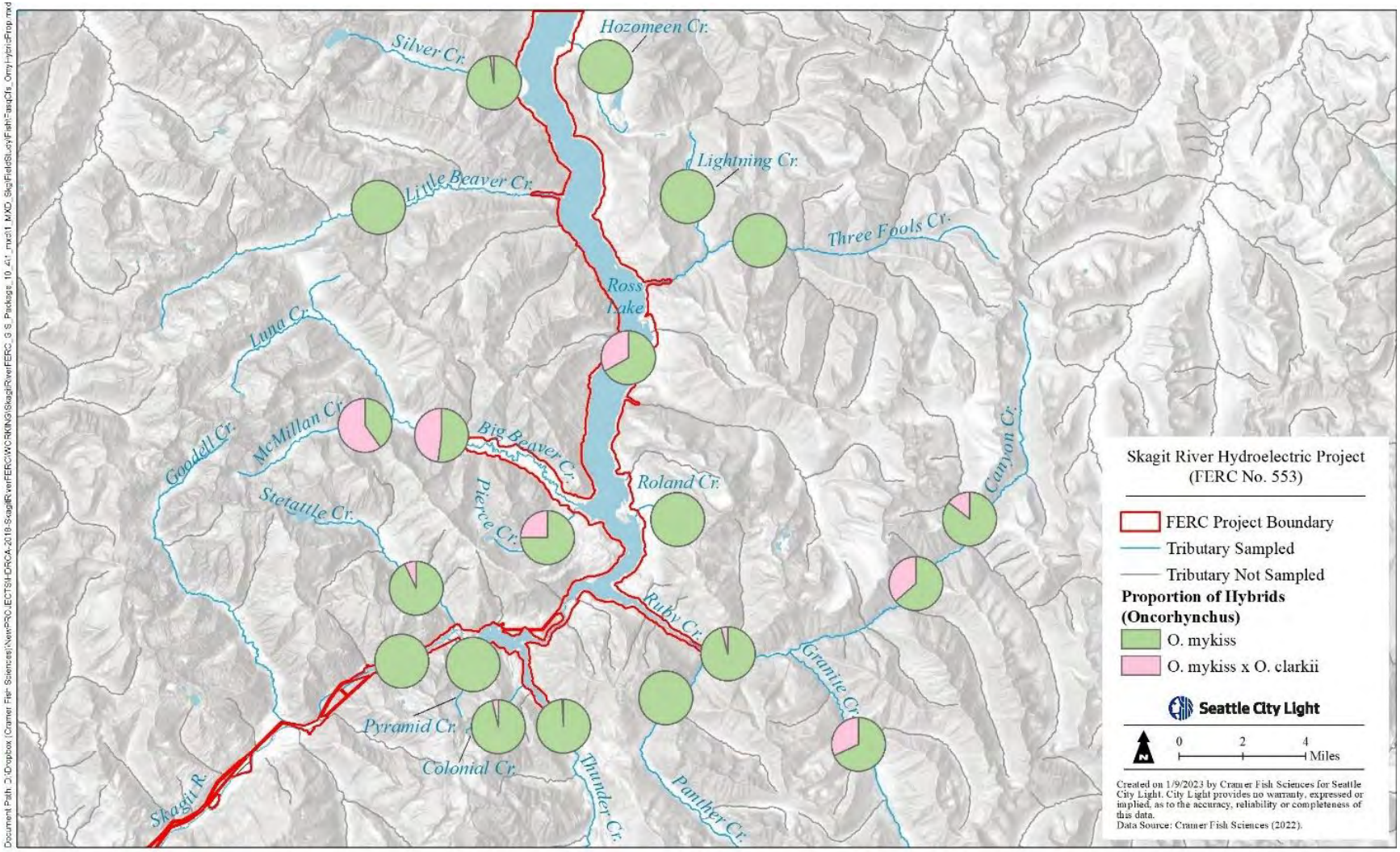


Figure 4. **Proportion of hybridized *Oncorhynchus* individuals observed by location. Hybrids were identified using three taxon-diagnostic SNPs. Thirteen percent of fish randomly collected within tributaries were identified as hybrids.**

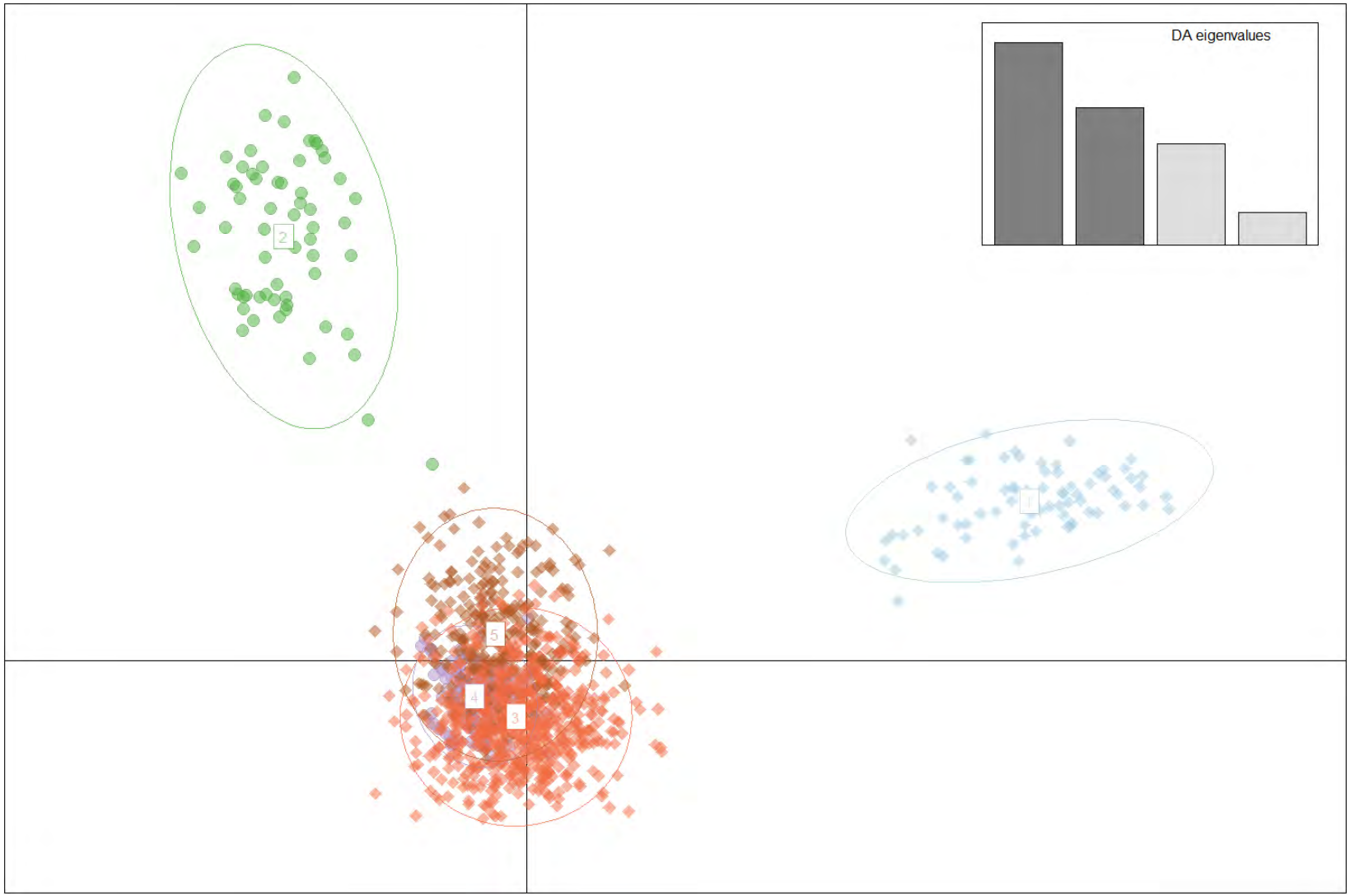


Figure 5. **Visual representation of DAPC analysis of Project *O. mykiss* for k=5 genetic clusters. Cluster-1 was Little Beaver Creek, cluster-2 was Pyramid Creek, cluster-4 was Three Fools Creek, and clusters-3 and -5 were remaining study fish. Note that cluster-4 is genetically distinct but is obscured on this two dimensional representation of underlying genetic variation data.**

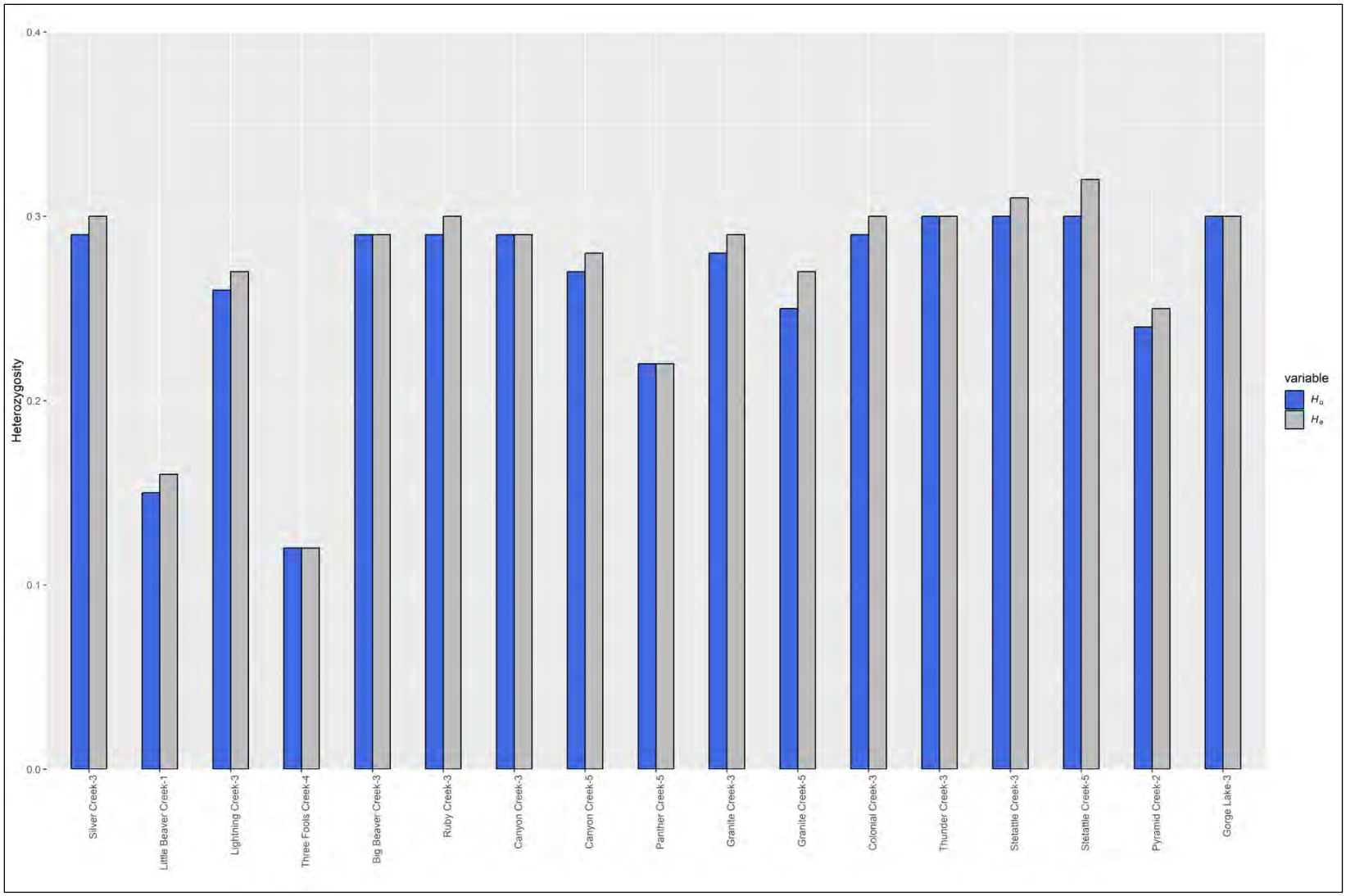


Figure 6. **Genetic diversity (mean observed and mean expected heterozygosity) for Project *O. mykiss* populations containing greater than 15 samples. Collections generally conformed to Hardy-Weinberg proportions, suggesting collections likely reflect the true population diversity and structure of *O. mykiss* in the Project vicinity.**

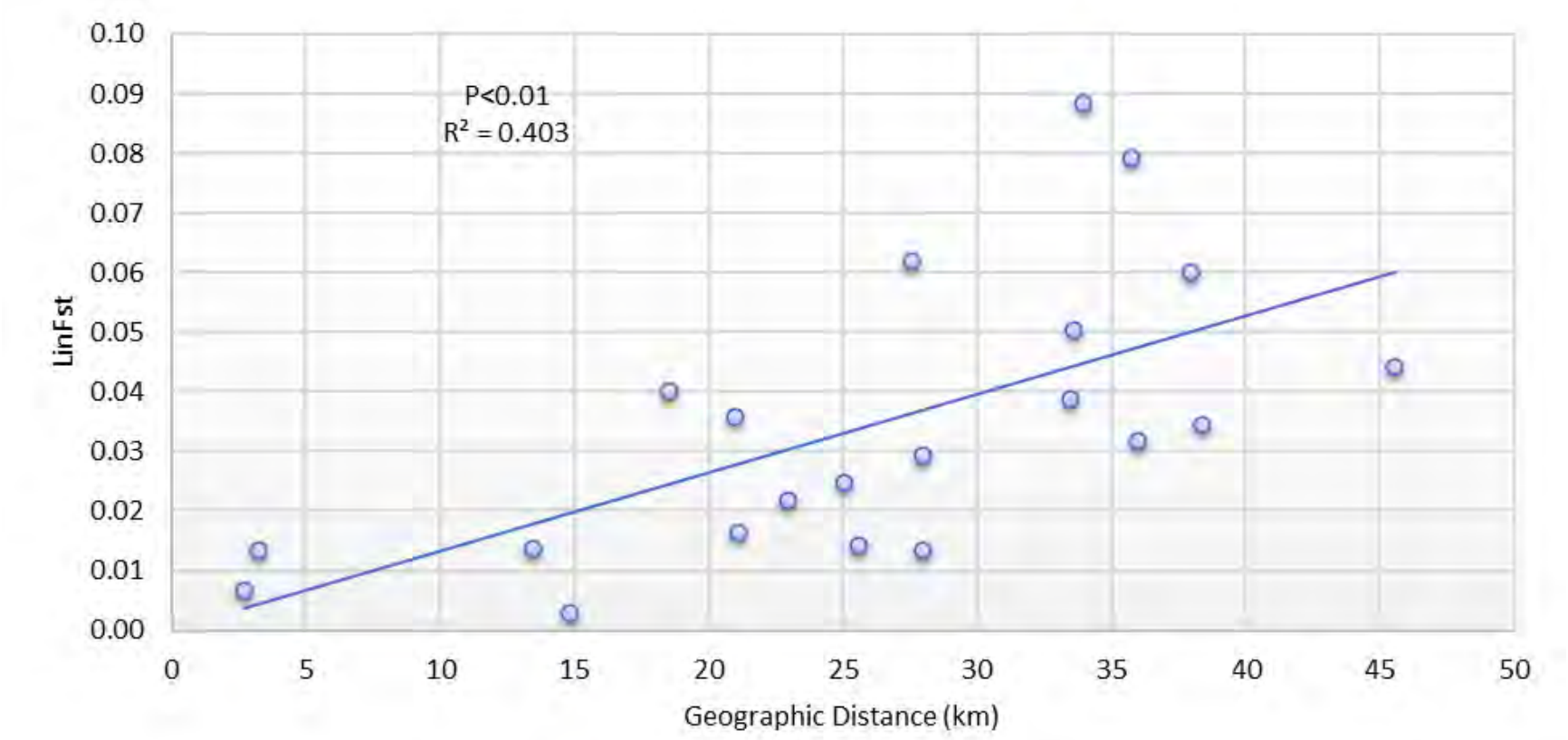


Figure 7. **Isolation by distance analyses for Dolly Varden in the study area assayed at 8 microsatellite loci. Note: Linear pairwise FST distances are plotted against pairwise geographical distance. The Mantel test suggests that 40 percent of the variability observed in the FST is explained by geographic distance.**

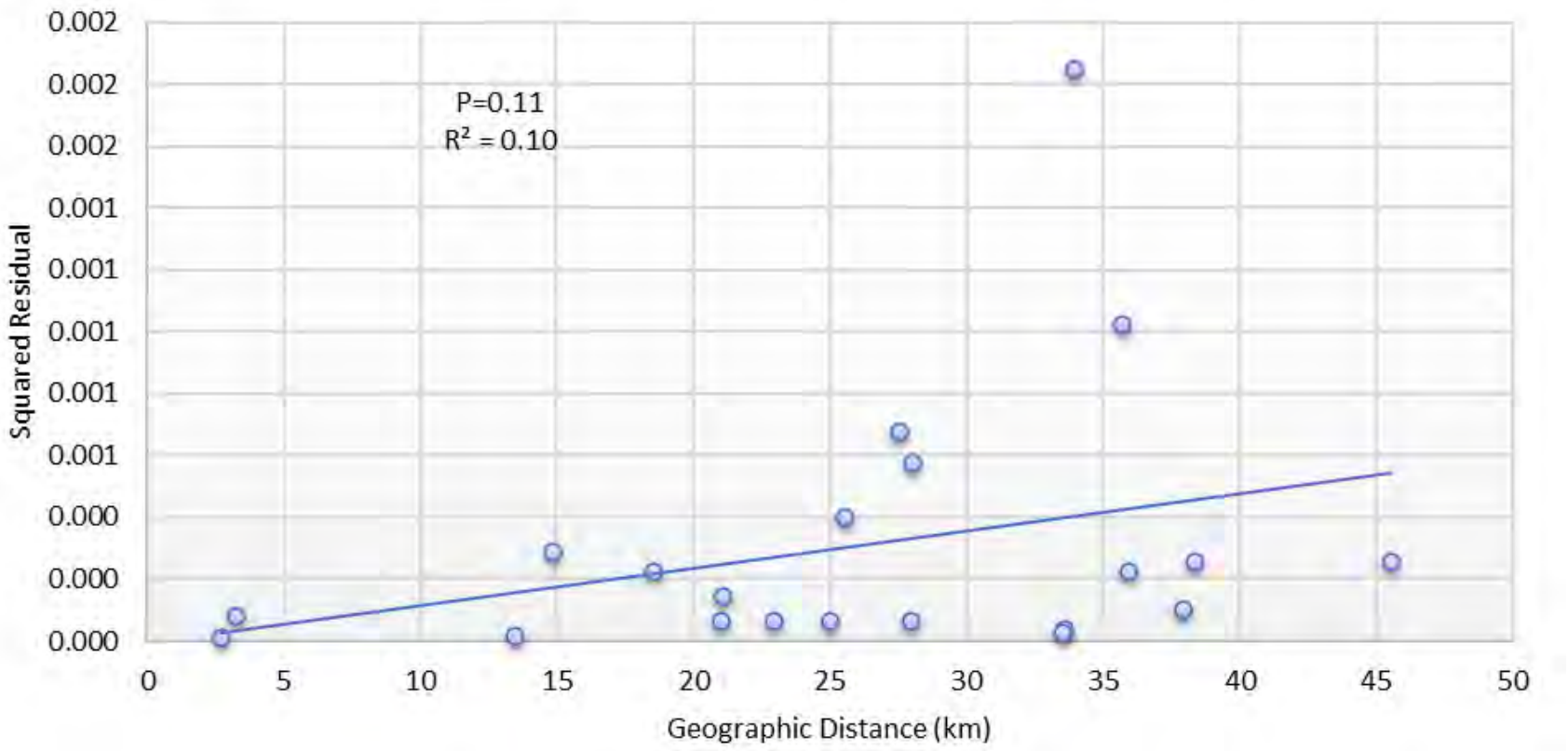


Figure 8. **Scatterplot of the squared residuals from the isolation-by-distance analysis in Dolly Varden using 8 microsatellites. Note: Although we observed a positive relationship (R2=0.10), the Mantel test was not statistically significant (P=0.11) and so the null hypothesis that Dolly Varden are not in migration-drift equilibrium cannot be rejected.**

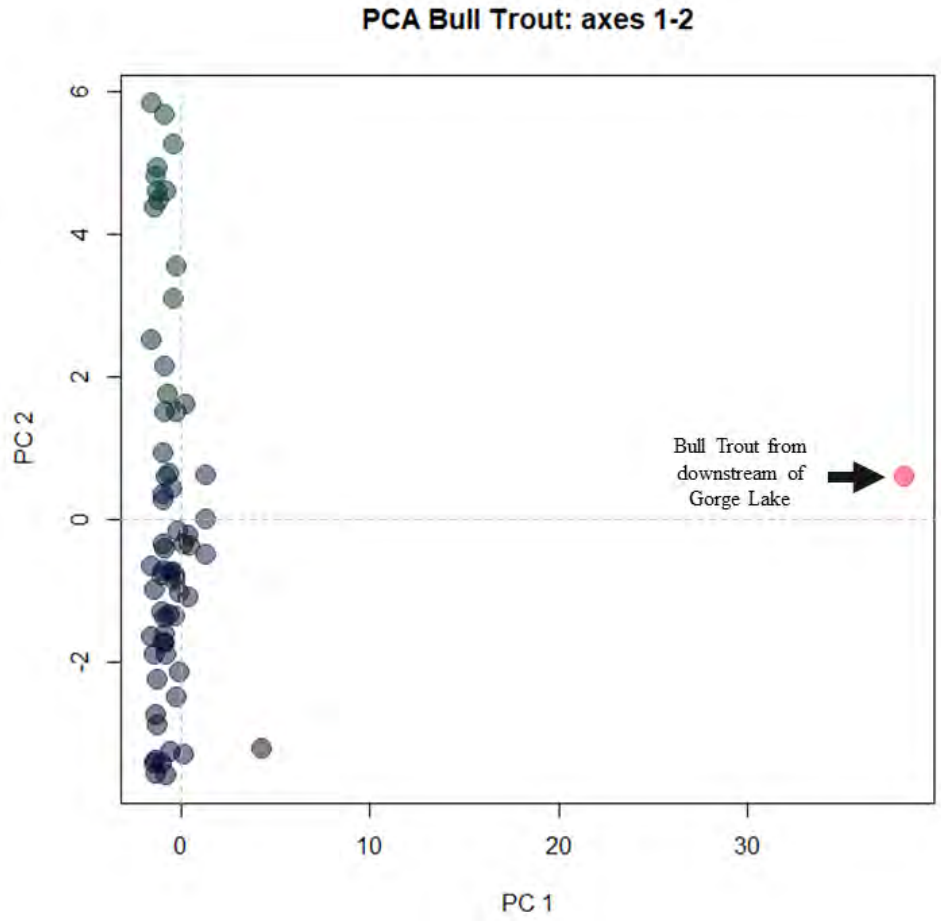


Figure 9. **Scatterplot of first two PCs based on genotypes of 66 Bull Trout sampled in Diablo, Gorge, and Ross lakes and downstream. PCs are shown on "Red, Green, Blue" (RGB) scale, which are indexed based on first three PCs. Greater PC distance = greater genetic distance and more color difference = more diversity. The plot highlights substantial differences in allele frequencies between Bull Trout from the Project Vicinity and downstream, with Bull Trout from the Project vicinity having substantially lower loadings for PC1. Higher "red" loadings highlight higher diversity downstream.**

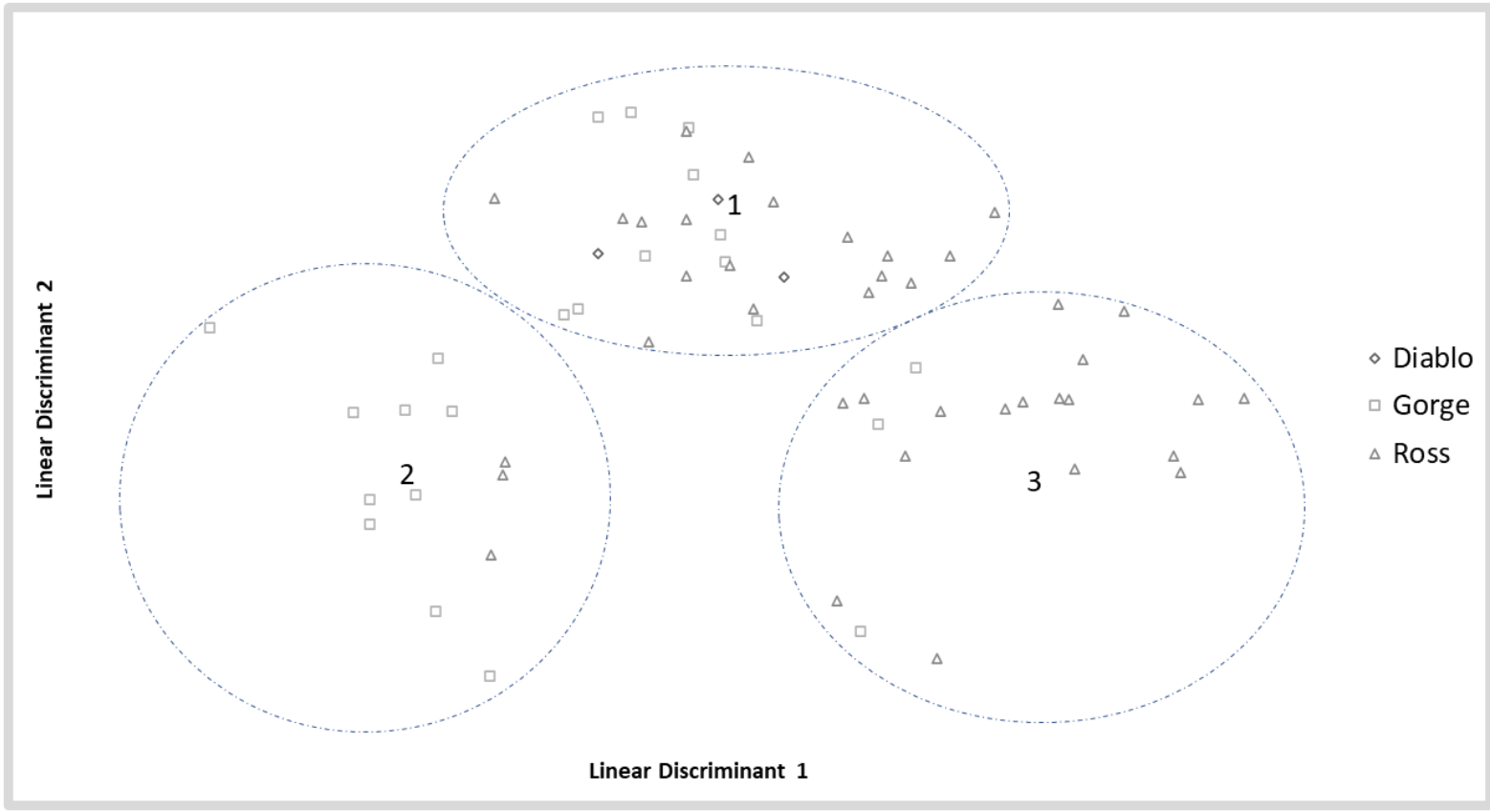


Figure 120. **Final DAPC analysis assuming k=3 for 65 Bull Trout sampled from the Project area during 2020, 2021, and 2022. The analysis used the first six PCs and two discriminant functions are plotted. The proportion of conserved variance was 0.54. Each point in the analysis represents an individual fish, and the shape of the point indicates which reservoir (Ross, Diablo, or Gorge) the fish was sampled from. The three ellipses encircle fish that assigned to each of the three inferred genetic clusters. The posterior probability of assignment to each cluster was 100 percent, indicating that there is strong genetic structuring present in the Project area. The model suggests each reservoir does not contain its own unique genetic population. Nevertheless, geographical based structure was apparent because fish from Ross had higher values of Discriminant Function 1, while fish from Gorge Lake had lower values, and fish from Diablo Lake fell in between. Sample sizes were small and so we cannot rule out the potential for isolation by distance.**

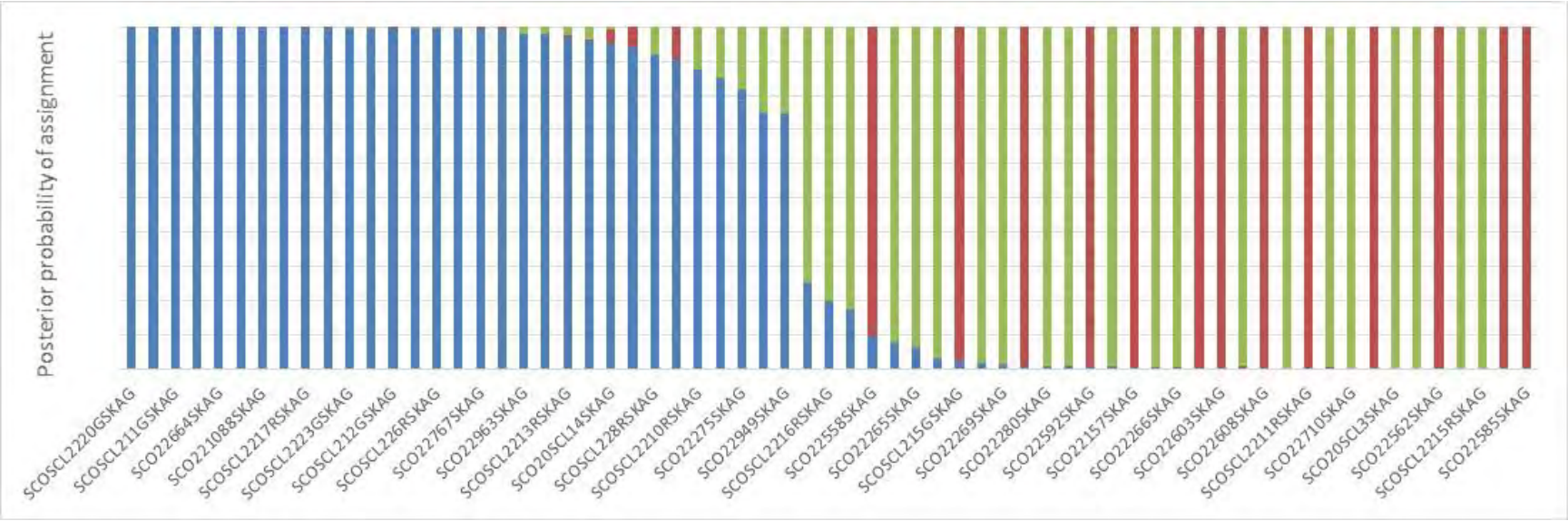


Figure 11. **Composition plot showing the posterior probability of assignment of Bull Trout to three inferred clusters. This plot is another way of looking at data presented in Figure 10. The individual fish ID is plotted along the bottom axis and each bar is an individual Bull Trout. The individuals are sorted (largest to smallest) by their posterior probability of assignment to each of three k. Fish with the acronym "SLC" were sampled at large from reservoirs and tended to be older fish than what was sampled via electrofishing in 2022. There could be age-related genetic structure but sample sizes were small (N=65).**

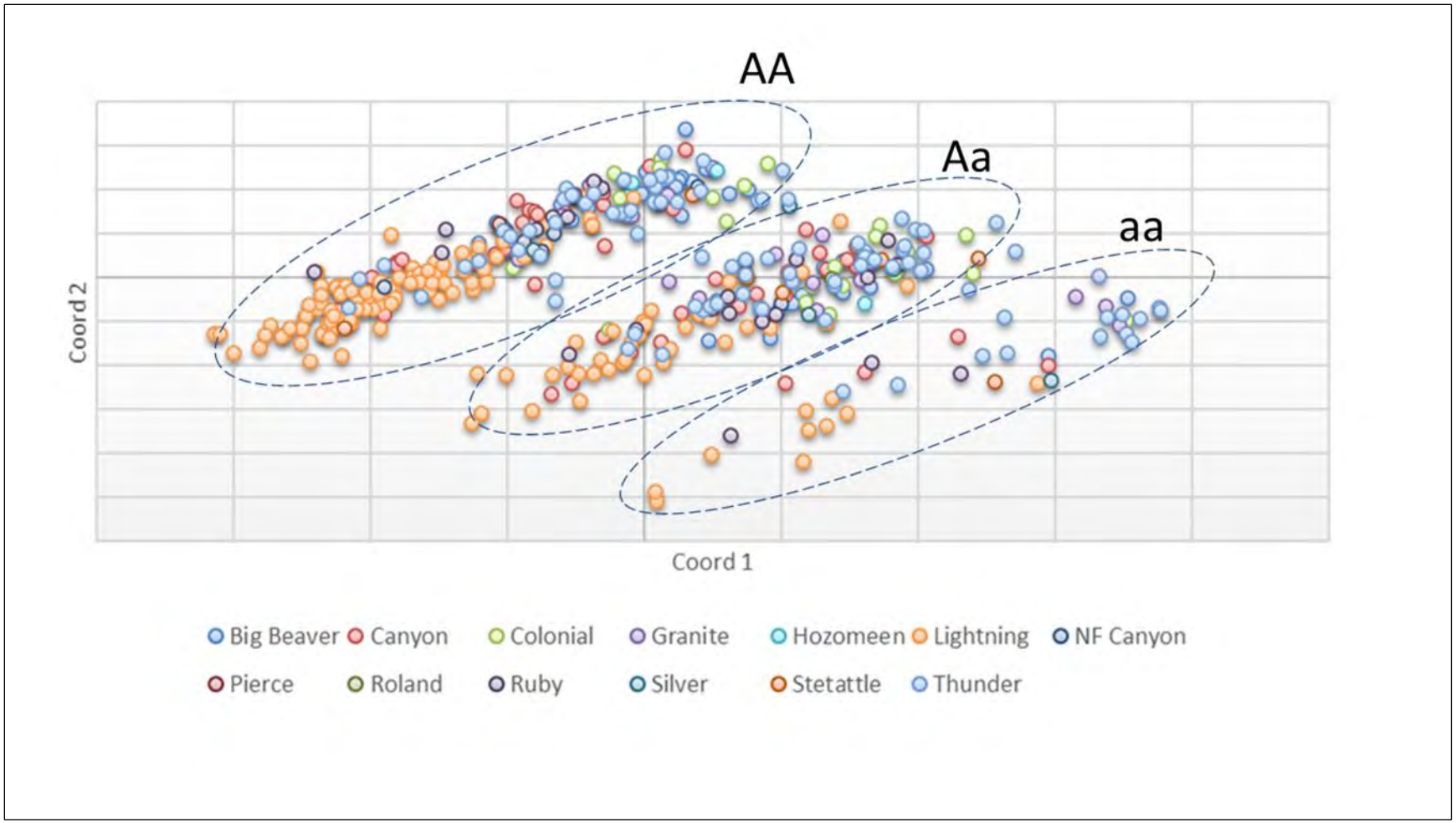


Figure 12 (1 of 2). Three groups are apparent. The three groups are in Hardy-Weinberg proportions across the Project area.

Figure 12. **PCoA of Dolly Varden genetic distance matrix. This page: Fish colored by tributary, showing three genetic groups and unexpected separation of fish sampled from within tributaries across the three groups. Next page: Fish colored by tributary and by membership to two inferred genetic groupings based on DAPC. The figure suggests a widespread contact zone between two genetic populations of Dolly Varden.**

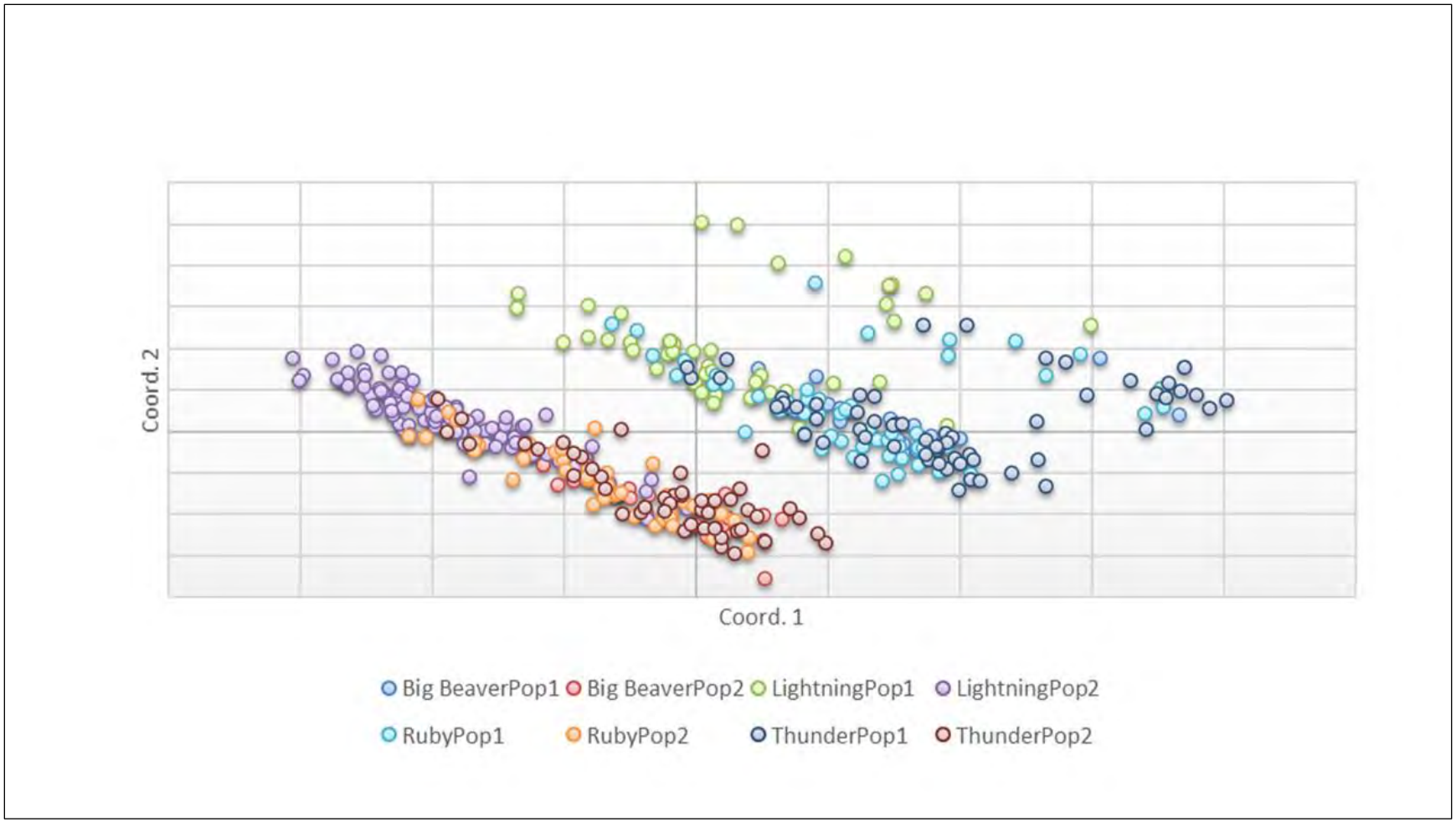


Figure 12 (2 of 2). Same as previous page but the projection is reversed due to a software glitch.

Figure 12. **PCoA of Dolly Varden genetic distance matrix. Previous page: Fish colored by tributary, showing three genetic groups and unexpected separation of fish sampled from within tributaries across the three groups. This page: Fish colored by tributary and by membership to two inferred genetic groupings based on DAPC. The figure suggests a widespread contact zone between two genetic populations of Dolly Varden.**

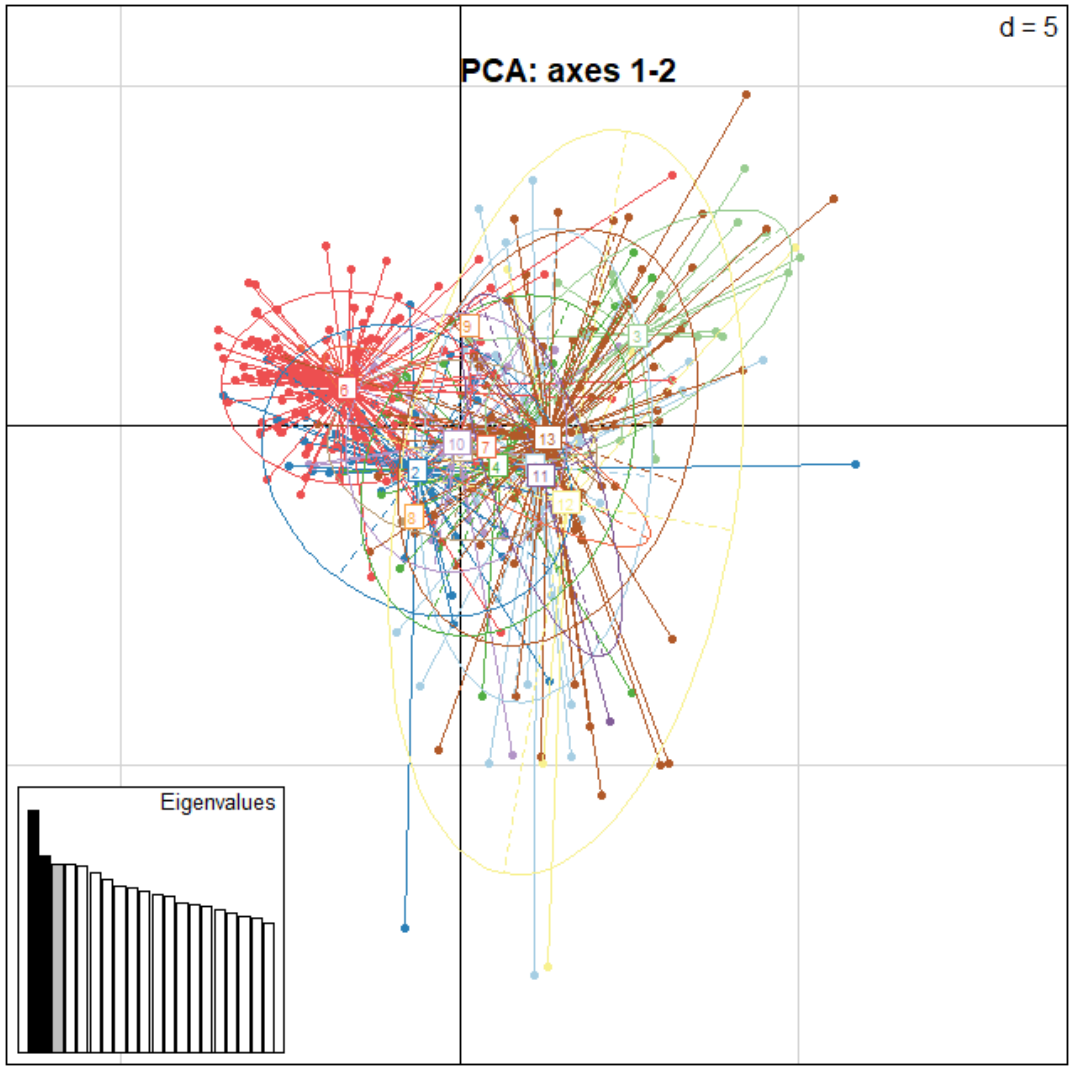


Figure 13. **Scatterplot of first two PCs based on genotypes at eight microsatellites in 413 Dolly Varden sampled in 13 collections from the Project vicinity. (1) Big Beaver, (2) Canyon, (3) Colonial, (4) Granite, (5) Hozomeen, (6) Lightning, (7) NF Canyon, (8) Pierce, (9) Roland, (10) Ruby, (11) Silver, (12) Stetattle, (13) Thunder creeks.**

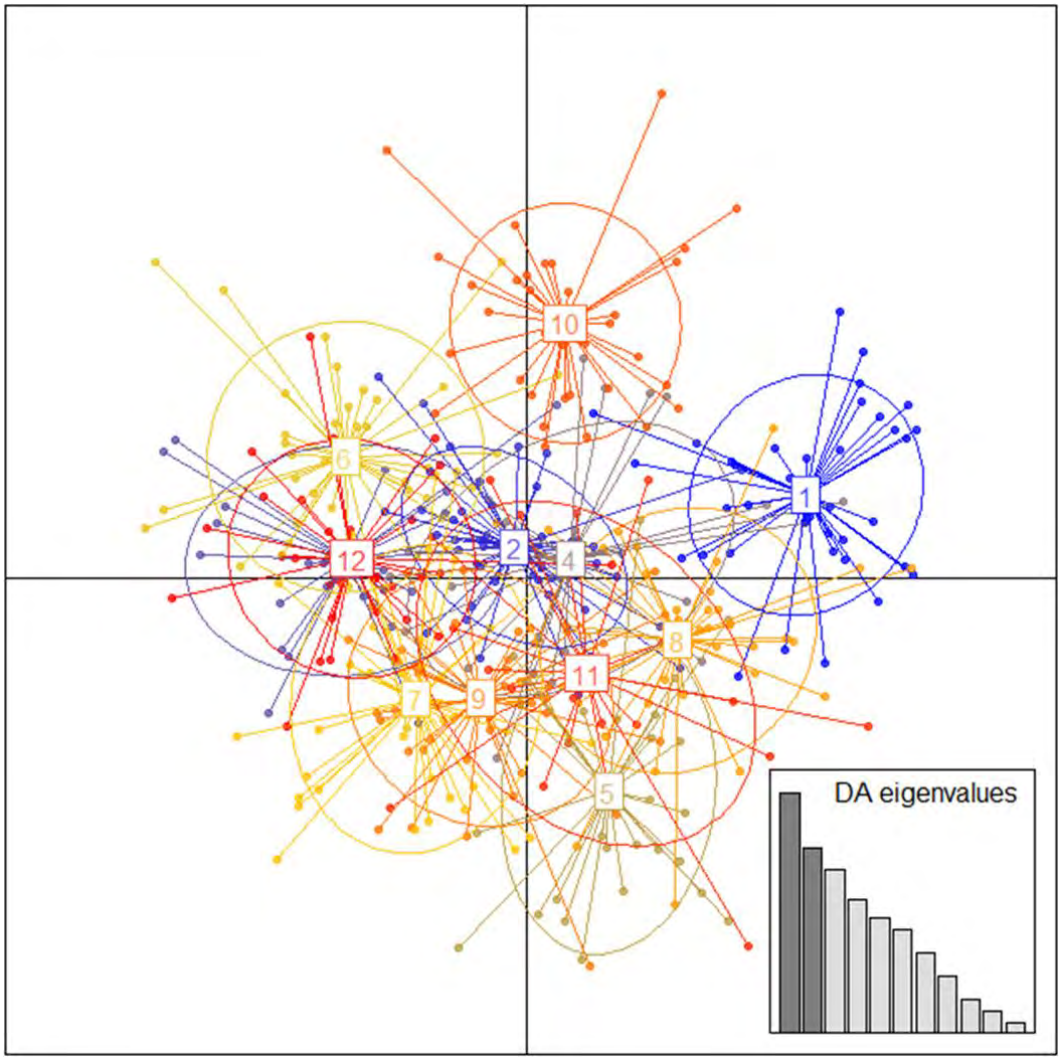


Figure 14 (1 of 2).

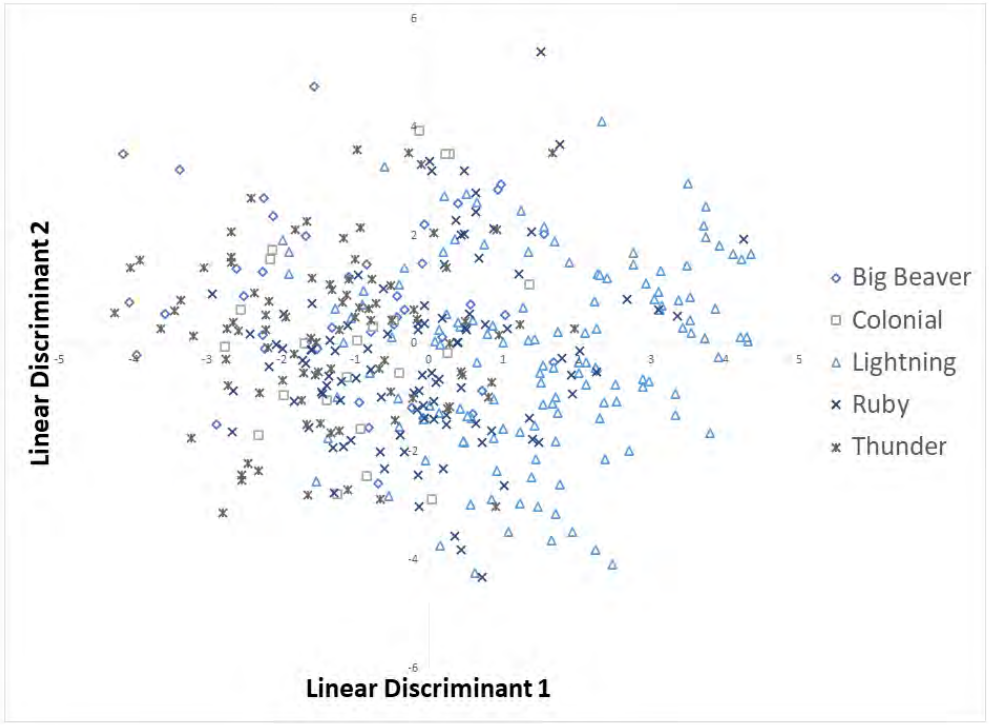


Figure 14 (2 of 2).

Figure 14. **Scatterplot of discriminant function scores for Dolly Varden at K=12; 40 PCs, 11 discriminant functions graph represents the individuals as dots and the groups as inertia ellipses. Note: Eigenvalues of the analysis are displayed in inset. Plots show the projection of the first two linear discriminants, where the x-axis represents the first linear discriminant, and the y-axis represents the second linear discriminant. The structure resembled a hierarchical pattern described in Jombart et al. (2010), which prompted exploratory analyses that resulted in discovery of the highly distinct genetic groupings discussed in the Results.**

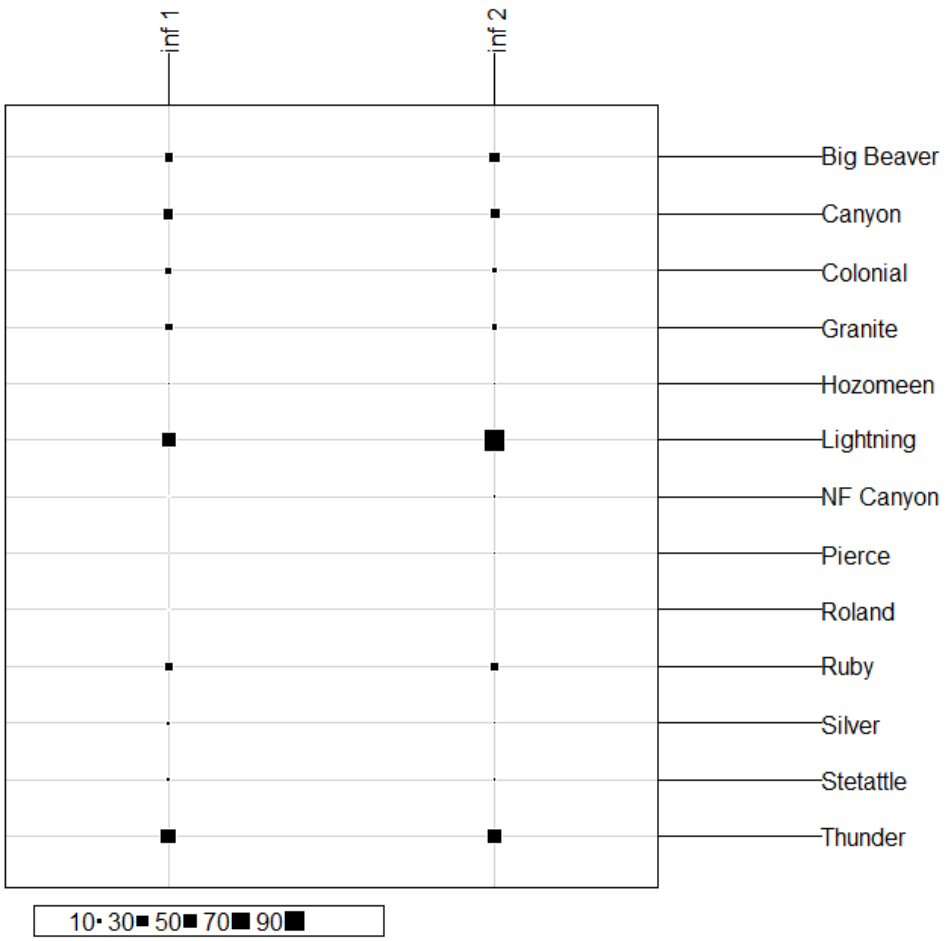


Figure 15. **K-means clustering of k=2 implemented to test the hypothesis that genetic structure of Dolly Varden in the study area driven by cryptic admixture. The columns across the top axis identify the two inferred clusters and the tributary names along the right-hand axis identify the sample locations. The size of the black box in each cell indicates how many fish from each inferred cluster were sampled in each tributary. The wide distribution of fish from both inferred clusters across nearly all tributaries suggests that there is a contact zone between the two groups of Dolly Varden.**

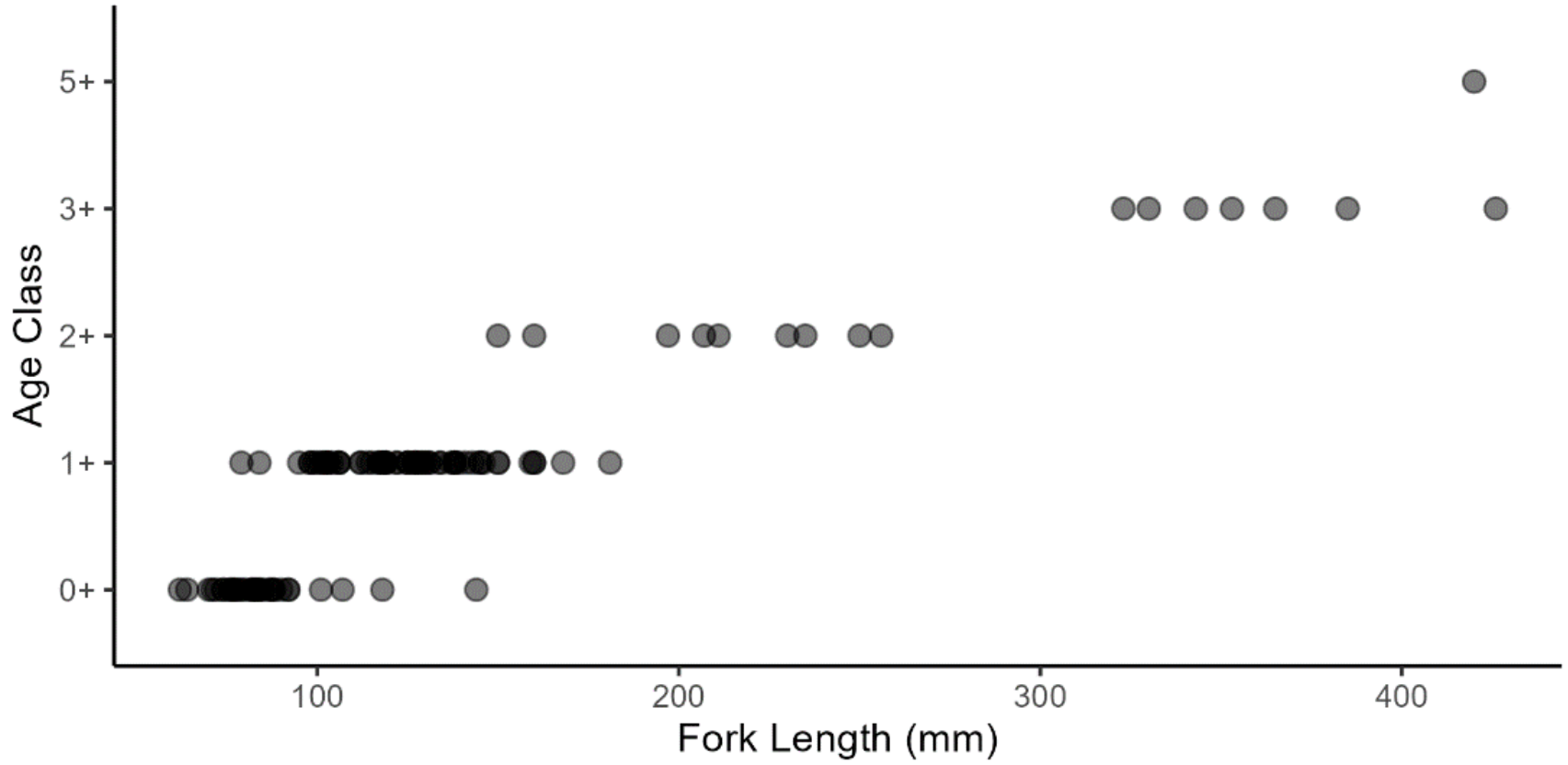


Figure 16. **Summary of age, based on scale analysis, and fork length (mm) of 110 *Salvelinus* captured in 2022.**

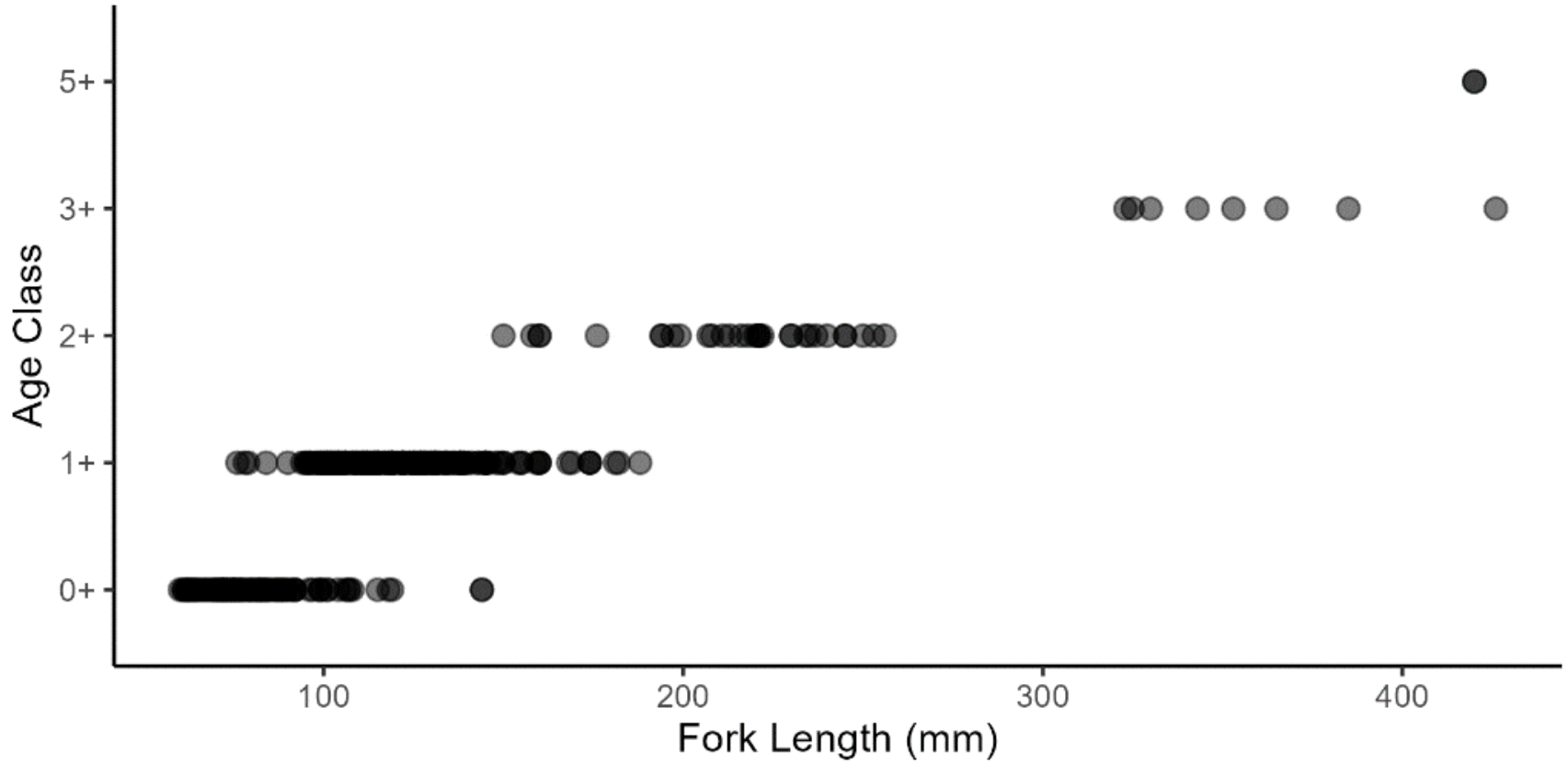


Figure 17. Summary of age, based on scale analysis and age-assignment, and fork length (mm) of 311 *Salvelinus*.

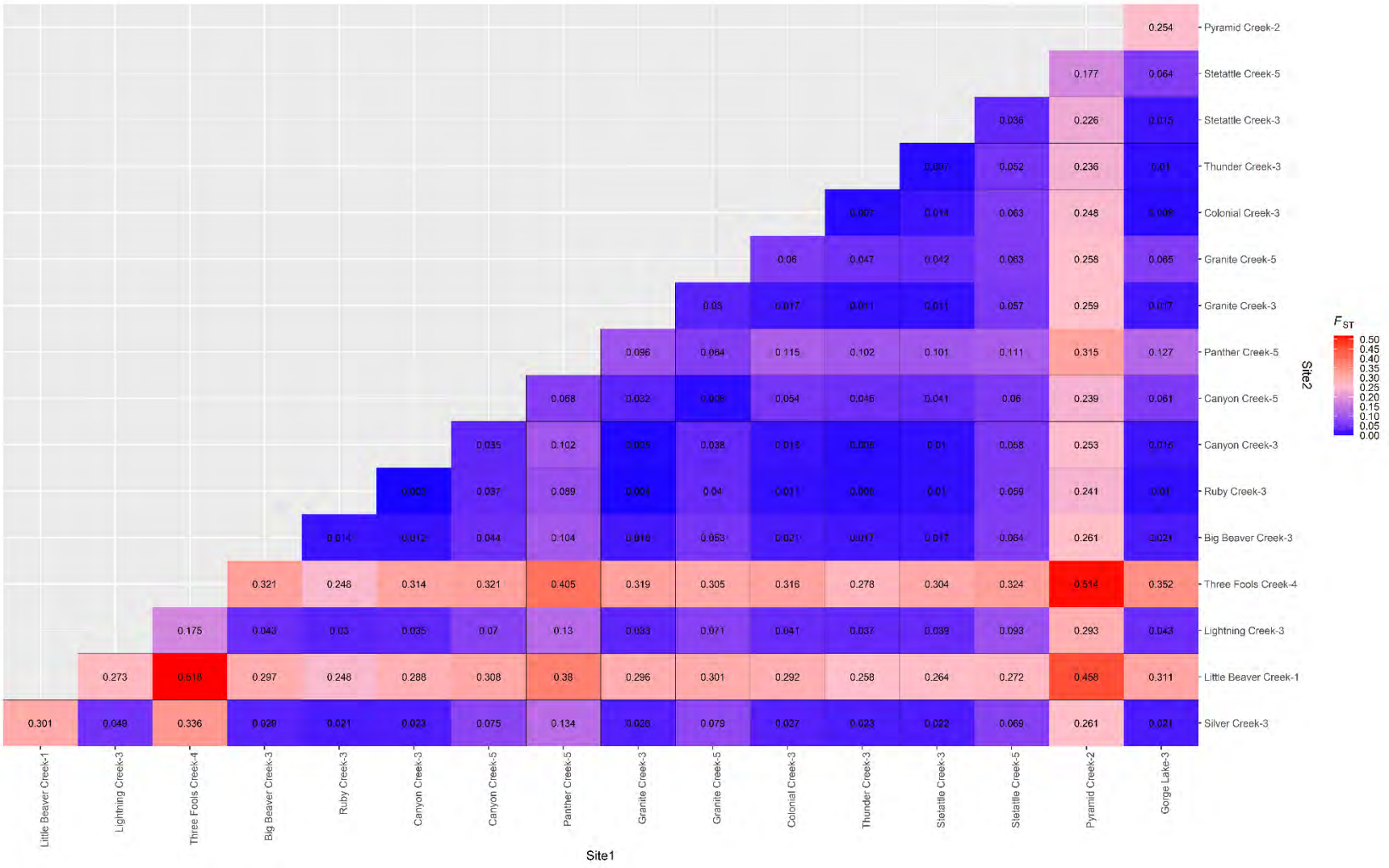


Figure 18. **Genetic distances between study area *O. mykiss* as represented by pairwise estimates of fixation index (F_{ST}). Populations analyzed contained greater than 15 samples. Note that estimates of $F_{ST} > 0.05$ could be considered substantial.**

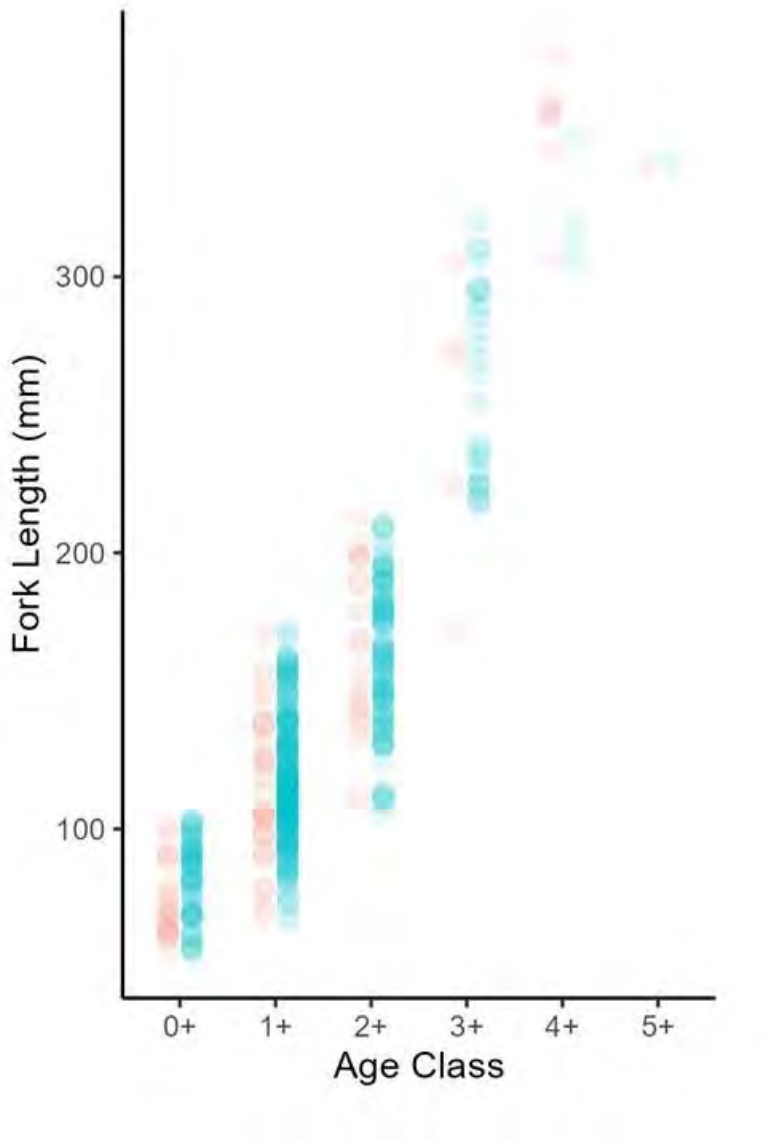


Figure 19. **Observed (red) and modeled (green) ages based on scale analysis and age-assignment, and fork length (mm) of 507 *Oncorhynchus*. A single fish, the smallest age-3+ category, was removed prior to modeling.**

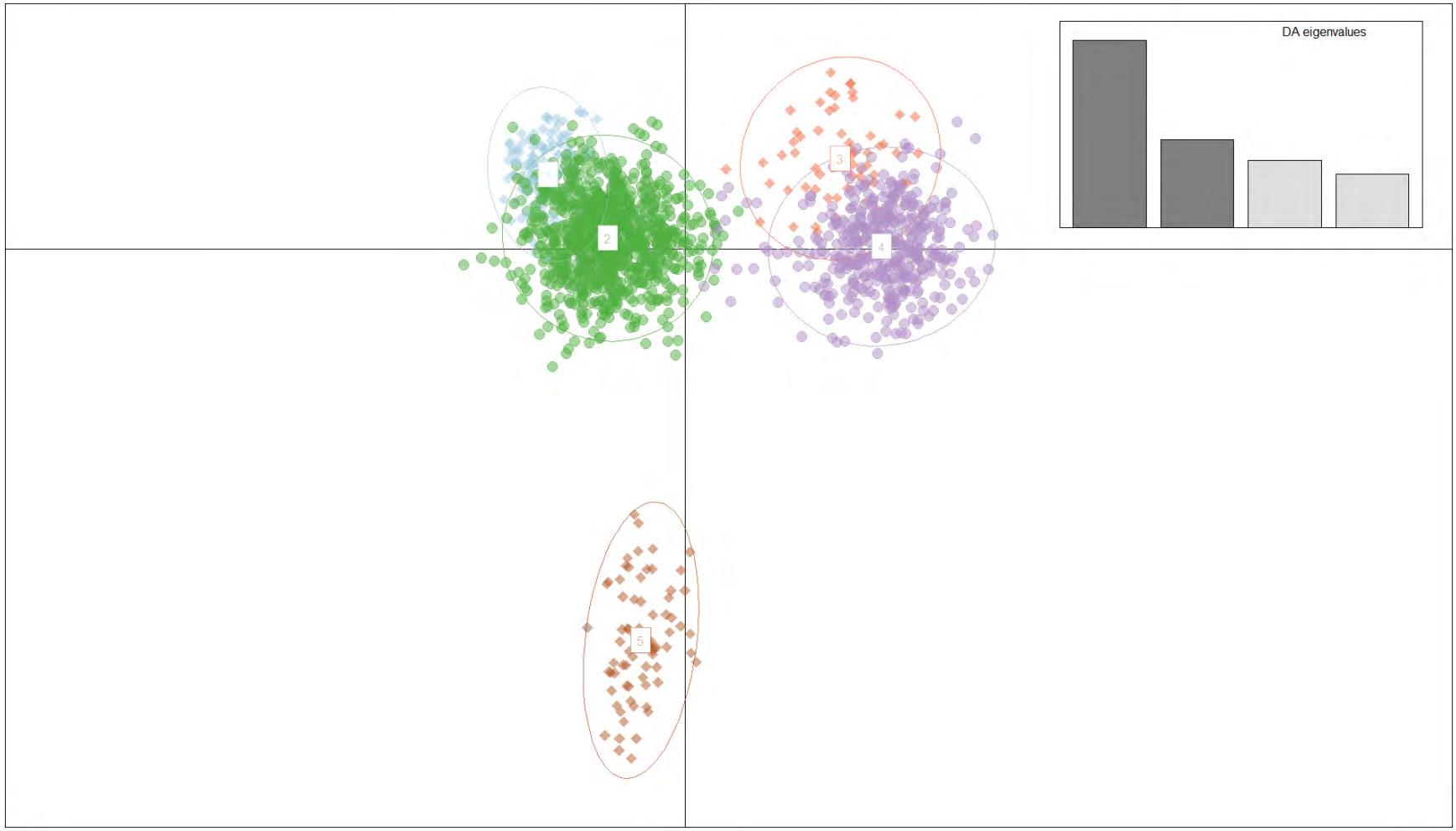


Figure 20. **Visual representation of DAPC analysis of above- and below-Project O. mykiss for k=5 genetic clusters. Cluster-1 is predominantly Three Fools Creek with some Lightning Creek, most study area fish resided in cluster-2, cluster-3 is Pyramid Creek, all but one fish from below Gorge Dam resided in cluster-4, and cluster-5 was Little Beaver Creek.**

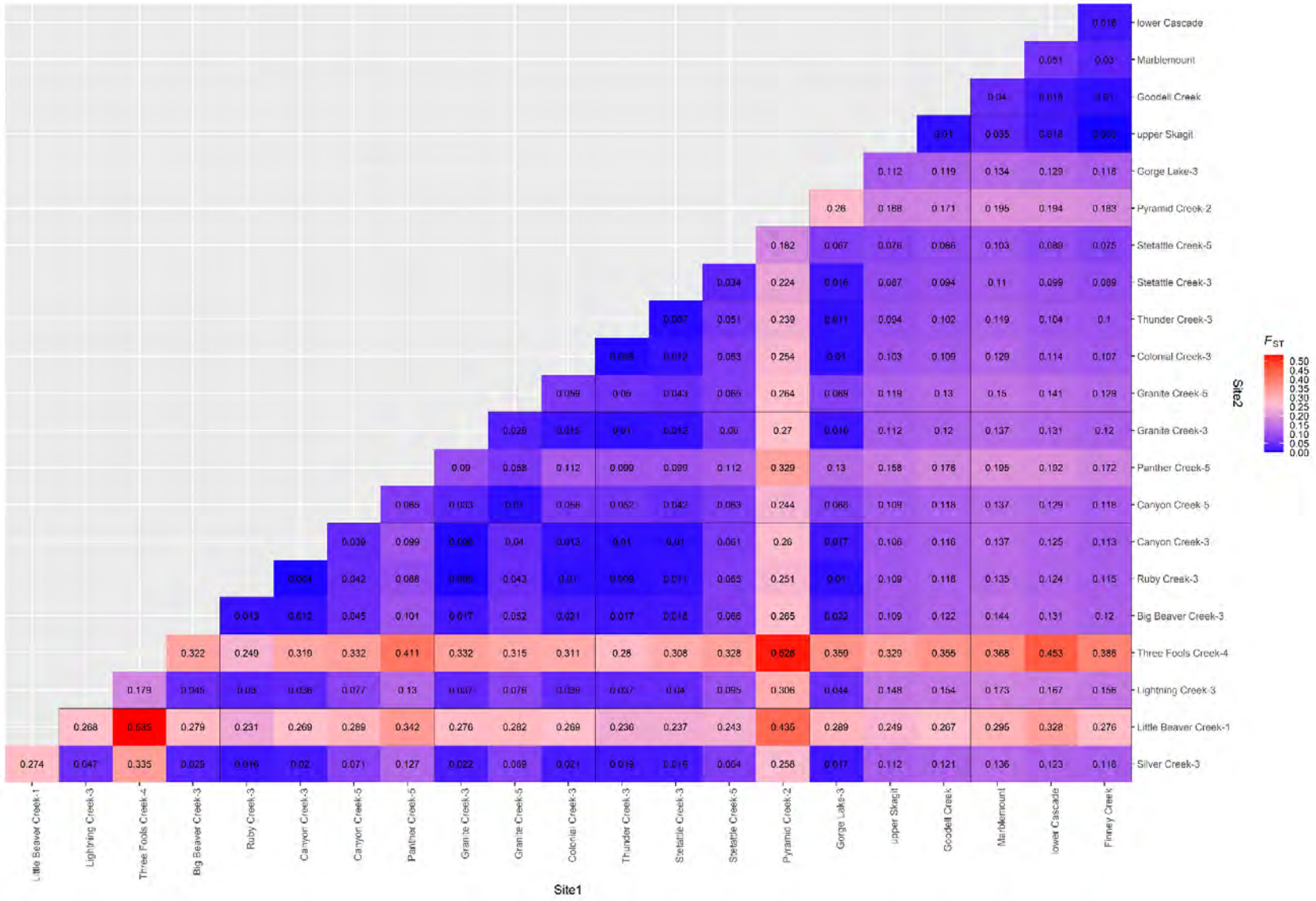


Figure 21. **Pairwise estimates of fixation index (F_{ST}) for above- and below-Project *O. mykiss* populations containing greater than 15 samples. Note that estimates of $F_{ST} > 0.05$ could be considered substantial.**

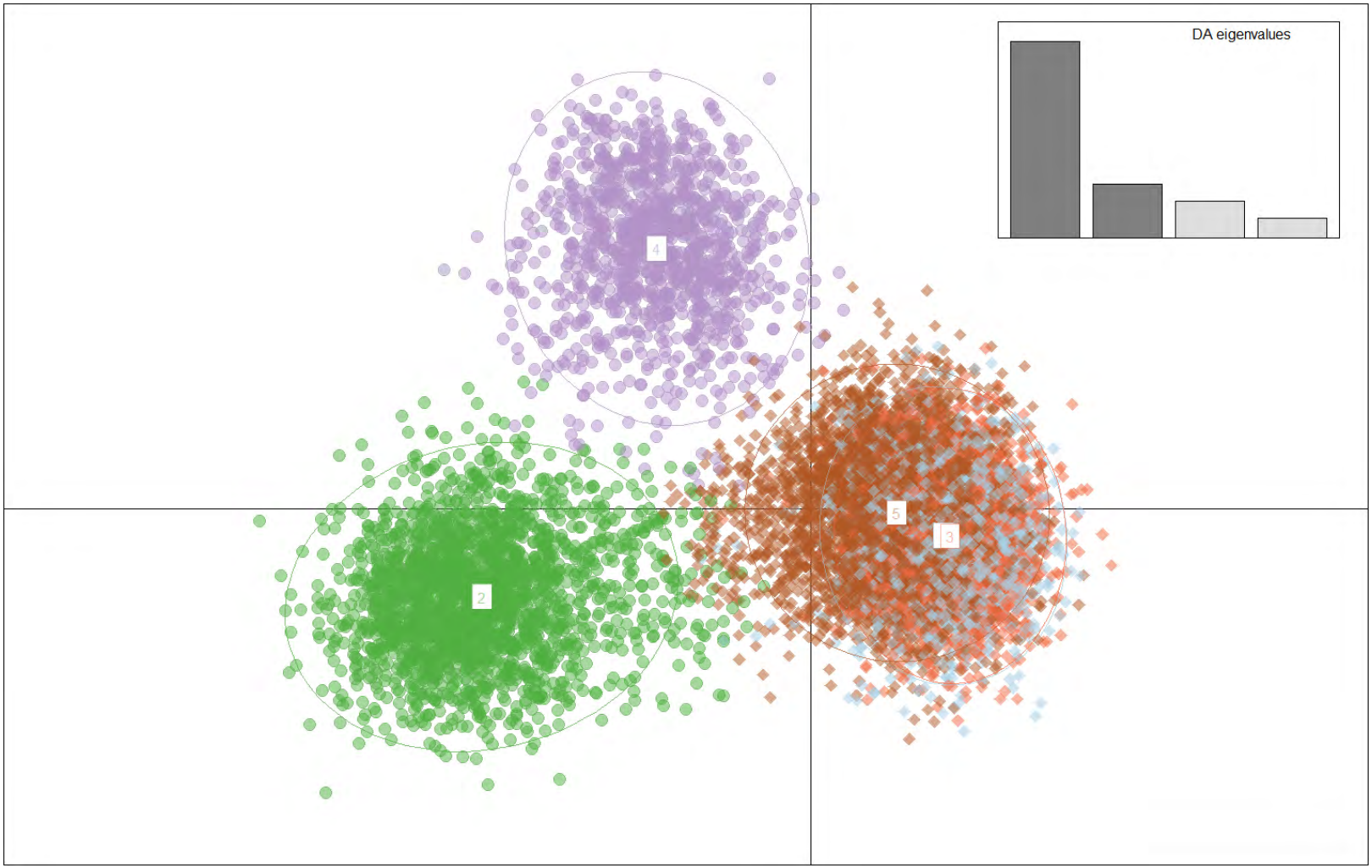


Figure 22. **Visual representation of DAPC analysis of regional *O. mykiss* dataset for k=5 genetic clusters. Note that Cluster-4 represents 29 (of 30) study area *O. mykiss* populations. Pyramid Creek population resided in Cluster-2 with other coastal *O. mykiss* populations. Clusters-1, -3, and -5 represented the redband subspecies of *O. mykiss* populations from East of the Cascade Mountains. The primary x-axis was driven by differences among subspecies of *O. mykiss*, with study area *O. mykiss* placed intermediately along this axis. The secondary y-axis is driven by differences among study area *O. mykiss* and all other *O. mykiss* in the regional dataset.**

RESERVOIR NATIVE FISH BASELINE GENETICS STUDY
TECHNICAL MEMORANDUM FOR EXPERT PANEL

ATTACHMENT A

SUMMARY OF ANALYSIS

Table A-1. Summary of k_2 for Bull Trout.

FID	k
SCO22767SKAG	2
SCO221088SKAG	2
SCO22963SKAG	2
SCO20SCL1SKAG	1
SCO20SCL26SKAG	2
SCOSCL211GSKAG	2
SCOSCL212GSKAG	2
SCOSCL213GSKAG	2
SCOSCL215GSKAG	1
SCOSCL2220GSKAG	2
SCOSCL2221GSKAG	2
SCOSCL2222GSKAG	2
SCOSCL2223GSKAG	2
SCO22208SKAG	1
SCO22209SKAG	2
SCO22211SKAG	1
SCO22558SKAG	1
SCO22560SKAG	2
SCO22562SKAG	1
SCO22579SKAG	2
SCO22585SKAG	1
SCO22592SKAG	1
SCO22603SKAG	1
SCO22604SKAG	1
SCO22608SKAG	1
SCO22664SKAG	2
SCO22948SKAG	2
SCO22949SKAG	2
SCO22253SKAG	2
SCO22156SKAG	1
SCO22157SKAG	1
SCO20SCL14SKAG	2
SCO20SCL15SKAG	2
SCO20SCL16SKAG	2
SCO20SCL22SKAG	2
SCO20SCL23SKAG	2
SCO20SCL2SKAG	2
SCO20SCL3SKAG	2
SCO20SCL4SKAG	2

SCO20SCL5SKAG	2
SCOSCL2210RSKAG	2
SCOSCL2211RSKAG	1
SCOSCL2213RSKAG	2
SCOSCL2215RSKAG	1
SCOSCL2216RSKAG	2
SCOSCL2217RSKAG	2
SCOSCL223RSKAG	2
SCOSCL224RSKAG	2
SCOSCL225RSKAG	2
SCOSCL226RSKAG	2
SCOSCL227RSKAG	2
SCOSCL228RSKAG	2
SCOSCL229RSKAG	2
SCO22258SKAG	2
SCO22259SKAG	2
SCO22265SKAG	2
SCO22266SKAG	2
SCO22269SKAG	2
SCO22274SKAG	2
SCO22275SKAG	2
SCO22280SKAG	2
SCO22469SKAG	2
SCO22506SKAG	1
SCO22710SKAG	2
SCO22722SKAG	2

Table A-2. Summary of k_3 for Bull Trout.

FID	k
SCO22767SKAG	1
SCO221088SKAG	1
SCO22963SKAG	1
SCO20SCL1SKAG	3
SCO20SCL26SKAG	1
SCOSCL211GSKAG	1
SCOSCL212GSKAG	1
SCOSCL213GSKAG	1
SCOSCL215GSKAG	2
SCOSCL2220GSKAG	1
SCOSCL2221GSKAG	3
SCOSCL2222GSKAG	3
SCOSCL2223GSKAG	1
SCO22208SKAG	2
SCO22209SKAG	1
SCO22211SKAG	2
SCO22558SKAG	2
SCO22560SKAG	1
SCO22562SKAG	2
SCO22579SKAG	1
SCO22585SKAG	2
SCO22592SKAG	2
SCO22603SKAG	2
SCO22604SKAG	2
SCO22608SKAG	2
SCO22664SKAG	1
SCO22948SKAG	1
SCO22949SKAG	1
SCO22253SKAG	3
SCO22156SKAG	2
SCO22157SKAG	2
SCO20SCL14SKAG	1
SCO20SCL15SKAG	3
SCO20SCL16SKAG	3
SCO20SCL22SKAG	1
SCO20SCL23SKAG	3
SCO20SCL2SKAG	3
SCO20SCL3SKAG	3
SCO20SCL4SKAG	1
SCO20SCL5SKAG	1

SCOSCL2210RSKAG	1
SCOSCL2211RSKAG	2
SCOSCL2213RSKAG	1
SCOSCL2215RSKAG	3
SCOSCL2216RSKAG	3
SCOSCL2217RSKAG	1
SCOSCL223RSKAG	1
SCOSCL224RSKAG	1
SCOSCL225RSKAG	1
SCOSCL226RSKAG	1
SCOSCL227RSKAG	1
SCOSCL228RSKAG	1
SCOSCL229RSKAG	3
SCO22258SKAG	3
SCO22259SKAG	1
SCO22265SKAG	3
SCO22266SKAG	3
SCO22269SKAG	3
SCO22274SKAG	3
SCO22275SKAG	1
SCO22280SKAG	3
SCO22469SKAG	3
SCO22506SKAG	3
SCO22710SKAG	3
SCO22722SKAG	1

Table A-3. Summary of k_4 for Bull Trout.

FID	k
SCO22767SKAG	4
SCO221088SKAG	4
SCO22963SKAG	2
SCO20SCL1SKAG	2
SCO20SCL26SKAG	4
SCOSCL211GSKAG	4
SCOSCL212GSKAG	4
SCOSCL213GSKAG	3
SCOSCL215GSKAG	1
SCOSCL2220GSKAG	4
SCOSCL2221GSKAG	2
SCOSCL2222GSKAG	3
SCOSCL2223GSKAG	4
SCO22208SKAG	1
SCO22209SKAG	4
SCO22211SKAG	1
SCO22558SKAG	1
SCO22560SKAG	4
SCO22562SKAG	1
SCO22579SKAG	4
SCO22585SKAG	1
SCO22592SKAG	1
SCO22603SKAG	1
SCO22604SKAG	1
SCO22608SKAG	1
SCO22664SKAG	4
SCO22948SKAG	4
SCO22949SKAG	3
SCO22253SKAG	2
SCO22156SKAG	2
SCO22157SKAG	1
SCO20SCL14SKAG	4
SCO20SCL15SKAG	2
SCO20SCL16SKAG	3
SCO20SCL22SKAG	3
SCO20SCL23SKAG	3
SCO20SCL2SKAG	3
SCO20SCL3SKAG	3
SCO20SCL4SKAG	4
SCO20SCL5SKAG	4

SCOSCL2210RSKAG	3
SCOSCL2211RSKAG	1
SCOSCL2213RSKAG	4
SCOSCL2215RSKAG	2
SCOSCL2216RSKAG	4
SCOSCL2217RSKAG	4
SCOSCL223RSKAG	4
SCOSCL224RSKAG	4
SCOSCL225RSKAG	4
SCOSCL226RSKAG	4
SCOSCL227RSKAG	4
SCOSCL228RSKAG	3
SCOSCL229RSKAG	2
SCO22258SKAG	2
SCO22259SKAG	2
SCO22265SKAG	3
SCO22266SKAG	2
SCO22269SKAG	3
SCO22274SKAG	2
SCO22275SKAG	3
SCO22280SKAG	3
SCO22469SKAG	3
SCO22506SKAG	2
SCO22710SKAG	3
SCO22722SKAG	4

Table A-4. Summary of k_12 for Dolly Varden.

x	
SCO22386SKAG	8
SCO22388SKAG	12
SCO22389SKAG	2
SCO22393SKAG	6
SCO22395SKAG	4
SCO22400SKAG	2
SCO22401SKAG	2
SCO22402SKAG	10
SCO22404SKAG	8
SCO22434SKAG	6
SCO22436SKAG	10
SCO22437SKAG	11
SCO22837SKAG	12
SCO22838SKAG	12
SCO22839SKAG	2
SCO22841SKAG	7
SCO22845SKAG	4
SCO22846SKAG	7
SCO22847SKAG	2
SCO22848SKAG	6
SCO22849SKAG	6
SCO22850SKAG	7
SCO22852SKAG	6
SCO22855SKAG	3
SCO22856SKAG	6
SCO22857SKAG	10
SCO22859SKAG	2
SCO22860SKAG	10
SCO22861SKAG	4
SCO22862SKAG	7
SCO22863SKAG	3
SCO22865SKAG	6
SCO22866SKAG	7
SCO22868SKAG	8
SCO22950SKAG	10
SCO22951SKAG	6
SCO22952SKAG	12
SCO22953SKAG	6
SCO22954SKAG	8
SCO22955SKAG	3

SCO22956SKAG	9
SCO22959SKAG	3
SCO22960SKAG	12
19PR0123	7
19PR0134	2
19PR0135	9
19PR0138	10
19PR0139	12
19PR0142	5
19PR0152	4
19PR0155	1
19PR0159	6
19PR0162	9
19PR0164	9
19PR0165	9
19PR0166	9
19PR0172	11
19PR0173	4
19PR0174	1
19QK0581	4
19QK0590	9
19QK0604	8
19QK0629	8
19QK0636	6
19QK0638	9
19QK0639	4
19QK0654	8
19QK0761	9
19QK0768	4
19QK0777	10
19QK1175	3
19QK1182	2
19QK1183	9
19QK1188	12
19QK1236	4
19QK1254	9
19QK1395	9
19QK1415	5
19QK1422	11
19QK1443	5
SCO22120SKAG	4
SCO22122SKAG	7

SCO22124SKAG	4
SCO22126SKAG	7
SCO22142SKAG	7
SCO22143SKAG	5
SCO22145SKAG	4
SCO22146SKAG	4
SCO22148SKAG	9
SCO22149SKAG	9
SCO22729SKAG	7
SCO22731SKAG	12
SCO22737SKAG	7
SCO22743SKAG	10
SCO22754SKAG	7
SCO22772SKAG	2
SCO22786SKAG	7
SCO22790SKAG	12
SCO22791SKAG	12
SCO22792SKAG	2
SCO22793SKAG	7
SCO22794SKAG	7
SCO22795SKAG	10
SCO22796SKAG	2
SCO22798SKAG	6
SCO22799SKAG	10
SCO22800SKAG	7
SCO22801SKAG	11
SCO22802SKAG	6
SCO22803SKAG	7
SCO22807SKAG	7
SCO22808SKAG	2
19PR0111	4
19PR0114	4
19PR0115	12
19PR0116	9
19PR0119	3
19PR0120	7
19PR0149	4
19PR0176	3
19PR0177	7
19PR0178	2
19PR0179	9
19PR0180	9

19PR0181	5
19PR0182	12
19QK1284	4
19QK1293	9
SCO22247SKAG	12
SCO22249SKAG	10
SCO22252SKAG	7
SCO22254SKAG	4
SCO22255SKAG	1
SCO22256SKAG	10
SCO22010SKAG	6
SCO22013SKAG	5
SCO22016SKAG	2
20NW0260	5
20NW0261	11
20NW0262	11
20NW0264	5
20NW0268	5
20NW0269	1
20NW0272	8
20NW0275	8
20NW0279	1
20NW0281	9
20NW0284	5
20NW0287	10
20NW0288	9
20NW0297	8
20NW0298	8
20NW0301	1
20NW0302	5
20NW0307	1
20NW0308	1
20NW0311	4
20NW0312	5
20NW0313	1
20NW0314	8
20NW0321	11
20NW0322	1
20NW0326	8
20NW0327	3
20NW0328	6
20NW0330	11

20NW0333	11
20NW0334	8
20NW0343	1
20NW0345	1
20NW0346	8
20NW0347	11
20NW0348	8
20NW0349	8
20NW0352	1
20NW0353	11
20NW0354	8
20NW0355	8
20NW0357	11
20NW0373	7
20NW0374	11
20NW0378	5
20NW0383	11
20NW0386	1
20NW0391	1
20NW0409	1
20NW0415	1
20NW0432	4
20NW0433	12
20NW0434	5
20NW0436	8
20NW0437	5
20NW0438	10
20NW0439	6
20NW0441	1
20NW0443	11
20NW0444	10
20NW0445	11
20NW0446	2
20NW0447	10
20NW0448	3
20NW0449	8
20NW0450	5
20NW0452	8
SCO22023SKAG	8
SCO22025SKAG	8
SCO22026SKAG	4
SCO22027SKAG	5

SCO22028SKAG	10
SCO22029SKAG	8
SCO22030SKAG	1
SCO22031SKAG	1
SCO22032SKAG	8
SCO22033SKAG	1
SCO22034SKAG	11
SCO22035SKAG	2
SCO22036SKAG	11
SCO22037SKAG	8
SCO22038SKAG	10
SCO22040SKAG	9
SCO22044SKAG	2
SCO22047SKAG	1
SCO22048SKAG	1
SCO22049SKAG	11
SCO22052SKAG	1
SCO22053SKAG	2
SCO22054SKAG	2
SCO22055SKAG	5
SCO22059SKAG	1
SCO22060SKAG	1
SCO22062SKAG	1
SCO22065SKAG	1
SCO22068SKAG	10
SCO22069SKAG	6
SCO22072SKAG	2
SCO22073SKAG	10
SCO22074SKAG	10
SCO22075SKAG	1
SCO22076SKAG	8
SCO22077SKAG	3
SCO22079SKAG	5
SCO22080SKAG	5
SCO22081SKAG	5
SCO22082SKAG	1
SCO22083SKAG	10
SCO22084SKAG	1
SCO22085SKAG	5
SCO22086SKAG	5
SCO22087SKAG	5
SCO22096SKAG	11

SCO22097SKAG	2
SCO22098SKAG	2
SCO22099SKAG	8
SCO22100SKAG	1
SCO22101SKAG	8
SCO22102SKAG	1
SCO22104SKAG	8
SCO22105SKAG	1
SCO22106SKAG	5
SCO22108SKAG	8
SCO22110SKAG	8
SCO22111SKAG	1
SCO22112SKAG	5
SCO22113SKAG	8
SCO22114SKAG	3
SCO22115SKAG	5
SCO22116SKAG	6
SCO22117SKAG	5
SCO22178SKAG	12
SCO22186SKAG	9
SCO22188SKAG	2
SCO22192SKAG	1
SCO22193SKAG	1
SCO22129SKAG	10
SCO22134SKAG	9
SCO22136SKAG	2
SCO22139SKAG	8
SCO22205SKAG	2
SCO22168SKAG	7
19PR0184	12
19PR0185	8
19PR0187	10
19PR0188	12
19PR0191	9
19PR0194	7
19PR0196	9
19PR0197	4
19PR0198	12
19PR0200	11
19PR0253	11
19PR0254	10
19PR0255	9

19QK0852	4
19QK0854	9
19QK0877	10
19QK0885	10
19QK0911	2
19QK0915	6
19QK1381	4
SCO22232SKAG	2
SCO22261SKAG	10
SCO22263SKAG	4
SCO22270SKAG	9
SCO22271SKAG	12
SCO22276SKAG	3
SCO22277SKAG	4
SCO22470SKAG	9
SCO22717SKAG	4
SCO22876SKAG	7
SCO22889SKAG	12
SCO22940SKAG	10
SCO22941SKAG	12
SCO22945SKAG	2
SCO22947SKAG	10
SCO22547SKAG	7
SCO22551SKAG	10
SCO22553SKAG	3
SCO22580SKAG	9
SCO22588SKAG	7
SCO22599SKAG	6
SCO22607SKAG	1
19PR0226	6
19PR0229	9
19PR0230	6
19PR0231	3
19PR0233	12
19PR0235	10
19PR0236	3
19PR0241	10
19PR0244	2
19PR0245	6
19PR0246	2
19PR0247	12
19QK1067	7

19QK1070	3
19QK1073	11
19QK1074	9
19QK1076	3
19QK1078	4
19QK1084	9
19QK1091	8
19QK1095	2
19QK1097	2
19QK1104	2
19QK1106	3
19QK1108	6
19QK1112	6
19QK1135	6
SCO221005SKAG	7
SCO221006SKAG	2
SCO221008SKAG	6
SCO221010SKAG	2
SCO221011SKAG	4
SCO221023SKAG	3
SCO221028SKAG	9
SCO221035SKAG	7
SCO221040SKAG	7
SCO221041SKAG	9
SCO221053SKAG	2
SCO221064SKAG	2
SCO221066SKAG	7
SCO221068SKAG	6
SCO221069SKAG	3
SCO221070SKAG	12
SCO221071SKAG	2
SCO221074SKAG	12
SCO221076SKAG	7
SCO221077SKAG	1
SCO221079SKAG	6
SCO221081SKAG	6
SCO221083SKAG	12
SCO221084SKAG	6
SCO221086SKAG	2
SCO221087SKAG	12
SCO221089SKAG	12
SCO221090SKAG	7

SCO221091SKAG	2
SCO221093SKAG	6
SCO221094SKAG	10
SCO221095SKAG	3
SCO221096SKAG	7
SCO221098SKAG	10
SCO221107SKAG	5
SCO221109SKAG	4
SCO221110SKAG	6
SCO221111SKAG	3
SCO221113SKAG	3
SCO221117SKAG	7
SCO221118SKAG	9
SCO221119SKAG	6
SCO221120SKAG	6
SCO221121SKAG	8
SCO221122SKAG	9
SCO221123SKAG	7
SCO221124SKAG	8
SCO221125SKAG	12
SCO221128SKAG	7
SCO221129SKAG	2
SCO221132SKAG	12
SCO221135SKAG	9
SCO22964SKAG	2
SCO22965SKAG	7
SCO22966SKAG	2
SCO22967SKAG	7
SCO22969SKAG	2
SCO22970SKAG	10
SCO22974SKAG	6
SCO22977SKAG	6
SCO22978SKAG	3
SCO22980SKAG	3
SCO22981SKAG	6
SCO22987SKAG	2
SCO22999SKAG	9

RESERVOIR NATIVE FISH BASELINE GENETICS STUDY
TECHNICAL MEMORANDUM FOR EXPERT PANEL

ATTACHMENT B

BAYESIAN INFORMATION CRITERION

Table B-1. BIC at k_2 for Bull Trout.

x	
K=1	102.26382
K=2	99.633745
K=3	98.431547
K=4	98.484856
K=5	98.856173
K=6	99.845205
K=7	100.6069
K=8	102.19042
K=9	103.36018
K=10	104.81224
K=11	106.79299
K=12	108.71607
K=13	110.00734
K=14	112.05278
K=15	113.70352

Table B-2. BIC at k_3 for Bull Trout.

x	
K=1	102.26382
K=2	99.633745
K=3	98.431547
K=4	98.508245
K=5	98.856173
K=6	99.783552
K=7	100.69873
K=8	101.92684
K=9	103.62176
K=10	104.9278
K=11	106.74335
K=12	108.70349
K=13	110.42847
K=14	112.25935
K=15	113.96841

Table B-3. BIC at k_4 for Bull Trout.

x	
K=1	102.26382
K=2	99.633745
K=3	98.431547
K=4	98.495041
K=5	98.856173
K=6	99.783552
K=7	100.81103
K=8	101.7053
K=9	103.23037
K=10	104.96759
K=11	107.0613
K=12	108.65454
K=13	110.55193
K=14	111.77638
K=15	113.87146

Table B-4. BIC at k_12 for Dolly Varden.

x	
K=1	399.2985508
K=2	375.356134
K=3	362.5325956
K=4	355.4499309
K=5	350.4349353
K=6	346.7476475
K=7	344.7336977
K=8	343.0362825
K=9	340.6576003
K=10	339.5929359
K=11	339.2610602
K=12	339.0177608
K=13	339.9377608
K=14	339.4581673
K=15	339.8042867
K=16	341.2383866
K=17	343.3631381
K=18	343.0067974
K=19	345.4740527
K=20	347.3199513

RESERVOIR NATIVE FISH BASELINE GENETICS STUDY
TECHNICAL MEMORANDUM FOR EXPERT PANEL

ATTACHMENT C

DISCRIMINANT ANALYSIS OF PRINCIPAL COMPONENTS

Figure C-1. DAPC analysis at k_2 for Bull Trout.

DAPC_k2
class: dapc
$call: dapc.genind(x = dat, pop = grp2$grp)
$n.pca: 20 first PCs of PCA used
$n.da: 1 discriminant functions saved
$var (proportion of conserved variance): 0.945
$eig (eigenvalues): 341.9

	vector	length	content
1	$eig	1	eigenvalues
2	$grp	65	prior group assignment
3	$prior	2	prior group probabilities
4	$assign	65	posterior group assignment
5	$pca.cent	73	centring vector of PCA
6	$pca.norm	73	scaling vector of PCA
7	$pca.eig	31	eigenvalues of PCA

	data.frame	nrow	ncol	content
1	$tab	65	20	retained PCs of PCA
2	$means	3	20	group means
3	$loadings	20	2	loadings of variables
4	$ind.coord	65	2	coordinates of individuals (principal components)
5	$grp.coord	3	2	coordinates of groups
6	$posterior	65	3	posterior membership probabilities
7	$pca.loadings	73	20	PCA loadings of original variables
8	$var.contr	73	2	contribution of original variables

Figure C-2. DAPC analysis at k_3 for Bull Trout.

dapc3
class: dapc
$call: dapc.genind(x = dat, pop = grp3$grp)
$n.pca: 20 first PCs of PCA used
$n.da: 2 discriminant functions saved
$var (proportion of conserved variance): 0.945
$eig (eigenvalues): 189.9 86.31

	Vector	length	content
1	$eig	2	eigenvalues
2	$grp	65	prior group assignment
3	$prior	3	prior group probabilities
4	$assign	65	posterior group assignment
5	$pca.cent	73	centring vector of PCA
6	$pca.norm	73	scaling vector of PCA
7	$pca.eig	31	eigenvalues of PCA

	data.frame	nrow	ncol	content
1	$tab	65	20	retained PCs of PCA
2	$means	3	20	group means
3	$loadings	20	2	loadings of variables
4	$ind.coord	65	2	coordinates of individuals (principal components)
5	$grp.coord	3	2	coordinates of groups
6	$posterior	65	3	posterior membership probabilities
7	$pca.loadings	73	20	PCA loadings of original variables
8	$var.contr	73	2	contribution of original variables

Figure C-3. DAPC analysis at k_4 for Bull Trout.

DAPC_4

class: dapc

$call: dapc.genind(x = dat, pop = grp$grp)

$n.pca: 15 first PCs of PCA used

$n.da: 3 discriminant functions saved

$var (proportion of conserved variance): 0.861

$eig (eigenvalues): 205.7, 54.3, 33.07

	vector	length	content
1	$eig	3	eigenvalues
2	$grp	65	prior group assignment
3	$prior	4	prior group probabilities
4	$assign	65	posterior group assignment
5	$pca.cent	73	centring vector of PCA
6	$pca.norm	73	scaling vector of PCA
7	$pca.eig	31	eigenvalues of PCA

	data.frame	nrow	ncol	content
1	$tab	65	15	retained PCs of PCA
2	$means	4	15	group means
3	$loadings	15	3	loadings of variables
4	$ind.coord	65	3	coordinates of individuals (principal components)
5	$grp.coord	4	3	coordinates of groups
6	$posterior	65	4	posterior membership probabilities
7	$pca.loadings	73	15	PCA loadings of original variables
8	$var.contr	73	3	contribution of original variables

Figure C-4. DAPC analysis at k_12 for Dolly Varden.

DACP_k12 DV

class: dapc

$call: dapc.genind(x = datDV, pop = grpDV$grp, n.pca = 50, n.da = 100)

$n.pca: 50 first PCs of PCA used

$n.da: 11 discriminant functions saved

$var (proportion of conserved variance): 0.979

$eig (eigenvalues): 109.1 82.01 73.8 61.6 54.71 ...

	vector	length	content
1	$eig	11	eigenvalues
2	$grp	413	prior group assignment
3	$prior	12	prior group probabilities
4	$assign	413	posterior group assignment
5	$pca.cent	102	centring vector of PCA
6	$pca.norm	102	scaling vector of PCA
7	$pca.eig	76	eigenvalues of PCA

	data.frame	nrow	ncol	content
1	$tab	413	50	retained PCs of PCA
2	$means	12	50	group means
3	$loadings	50	11	loadings of variables
4	$ind.coord	413	11	coordinates of individuals (principal components)
5	$grp.coord	12	11	coordinates of groups
6	$posterior	413	12	posterior membership probabilities
7	$pca.loadings	102	50	PCA loadings of original variables
8	$var.contr	102	11	contribution of original variables

Figure C-5. DAPC_ Posterior at k_3 for Bull Trout.

	1	2	3
SCO22767SKAG	0.9947175	0.005108	0.0001745
SCO221088SKAG	0.9995259	4.36E-05	0.0004305
SCO22963SKAG	0.9814204	0.0001683	0.0184112
SCO20SCL1SKAG	0.0002169	0.0032978	0.9964853
SCO20SCL26SKAG	0.9962899	0.0003664	0.0033437
SCOSCL211GSKAG	0.9999664	3.09E-05	2.67E-06
SCOSCL212GSKAG	0.9976194	0.0018708	0.0005099
SCOSCL213GSKAG	0.9613416	0.0015193	0.037139
SCOSCL215GSKAG	0.0198203	0.9801607	1.90E-05
SCOSCL2220GSKAG	0.9999828	1.24E-05	4.80E-06
SCOSCL2221GSKAG	0.0741189	0.000351	0.9255301
SCOSCL2222GSKAG	0.1723346	4.26E-05	0.8276227
SCOSCL2223GSKAG	0.998523	0.000154	0.001323
SCO22208SKAG	8.41E-06	0.9999891	2.46E-06
SCO22209SKAG	0.9418515	0.0575779	0.0005706
SCO22211SKAG	0.0001155	0.9998842	2.88E-07
SCO22558SKAG	0.0947729	0.9052136	1.34E-05
SCO22560SKAG	0.9999685	5.76E-06	2.57E-05
SCO22562SKAG	4.59E-05	0.9999538	2.42E-07
SCO22579SKAG	0.9040323	0.0954862	0.0004815
SCO22585SKAG	2.54E-06	0.9999799	1.76E-05
SCO22592SKAG	0.0073527	0.9926448	2.47E-06
SCO22603SKAG	0.0018393	0.9981607	7.87E-10
SCO22604SKAG	0.0020574	0.9979423	2.48E-07
SCO22608SKAG	0.0004044	0.9995934	2.23E-06
SCO22664SKAG	0.9998519	2.97E-05	0.0001184
SCO22948SKAG	0.9984946	1.35E-05	0.0014919
SCO22949SKAG	0.7479517	1.35E-06	0.2520469
SCO22253SKAG	0.0317156	2.21E-05	0.9682623
SCO22156SKAG	0.0104983	0.9893578	0.0001439
SCO22157SKAG	0.0061916	0.9936884	0.00012
SCO20SCL14SKAG	0.952247	0.0404291	0.0073239
SCO20SCL15SKAG	0.01738	0.0001439	0.9824762
SCO20SCL16SKAG	0.007489	2.89E-08	0.9925109
SCO20SCL22SKAG	0.7518492	8.61E-06	0.2481421
SCO20SCL23SKAG	3.01E-05	8.61E-10	0.9999699
SCO20SCL2SKAG	0.0036894	1.07E-07	0.9963105
SCO20SCL3SKAG	7.57E-05	3.16E-11	0.9999243
SCO20SCL4SKAG	0.9991278	0.0007701	0.0001021

	1	2	3
SCO20SCL5SKAG	0.9955191	0.000381	0.0040999
SCOSCL2210RSKAG	0.8759895	1.56E-05	0.123995
SCOSCL2211RSKAG	0.0002409	0.9997176	4.15E-05
SCOSCL2213RSKAG	0.9741224	0.0010639	0.0248138
SCOSCL2215RSKAG	1.88E-05	0.0001388	0.9998425
SCOSCL2216RSKAG	0.1984185	0.000485	0.8010965
SCOSCL2217RSKAG	0.9992594	0.0005706	0.00017
SCOSCL223RSKAG	0.993243	0.0067535	3.48E-06
SCOSCL224RSKAG	0.9997707	4.41E-06	0.0002249
SCOSCL225RSKAG	0.9999663	6.94E-06	2.68E-05
SCOSCL226RSKAG	0.9962463	0.0015024	0.0022513
SCOSCL227RSKAG	0.999418	0.0001821	0.0003999
SCOSCL228RSKAG	0.9183149	6.89E-06	0.0816782
SCOSCL229RSKAG	0.2514128	0.0012081	0.7473791
SCO22258SKAG	5.75E-05	1.05E-09	0.9999425
SCO22259SKAG	0.848722	3.76E-05	0.1512404
SCO22265SKAG	0.0637052	5.87E-08	0.9362947
SCO22266SKAG	0.002953	7.17E-08	0.997047
SCO22269SKAG	0.0136998	3.53E-09	0.9863002
SCO22274SKAG	0.0003185	9.01E-08	0.9996814
SCO22275SKAG	0.8169853	1.02E-07	0.1830146
SCO22280SKAG	0.0085946	1.24E-06	0.9914042
SCO22469SKAG	0.0073173	5.45E-07	0.9926822
SCO22506SKAG	0.0008938	0.0075528	0.9915534
SCO22710SKAG	0.0001952	2.44E-10	0.9998048
SCO22722SKAG	0.9791629	8.90E-06	0.0208282

Figure C-6. DAPC_ Posterior at k_12 for Dolly Varden.

	1	2	3	4	5	6	7	8	9	10	11	12
SCO22386SKAG	2.28E-06	0.1049222	0.1277036	0.0156598	2.44E-05	0.0047083	0.0001416	0.7442038	0.00111153	0.00011142	8.64E-06	0.00140228
SCO22388SKAG	2.26E-11	5.37E-08	1.25E-07	4.14E-12	2.61E-09	0.6940927	9.17E-08	2.20E-10	2.02E-08	1.87E-12	3.71E-09	0.30590702
SCO22389SKAG	4.25E-05	0.8876558	1.86E-05	7.66E-06	3.83E-07	0.0002015	0.0010251	1.19E-06	0.0009623	0.00016542	2.79E-08	0.10991953
SCO22393SKAG	5.64E-09	0.0004748	1.51E-07	0.0024192	7.25E-08	0.984666	8.17E-06	5.45E-08	2.54E-06	1.00E-08	1.15E-10	0.01242903
SCO22395SKAG	0.4998613	0.3211448	2.60E-06	0.177969	1.30E-07	9.98E-05	3.23E-05	1.02E-05	0.00064997	0.0002145	3.30E-06	1.20E-05
SCO22400SKAG	2.27E-05	0.3959531	3.37E-07	0.0192423	0.5781114	1.36E-05	5.51E-05	4.25E-05	0.00624865	2.50E-06	4.87E-06	0.000303
SCO22401SKAG	0.0001333	0.9725235	0.0008106	0.0086159	3.03E-06	0.0085081	0.0002253	0.0013916	0.00551533	0.00145495	9.23E-06	0.00080917
SCO22402SKAG	3.22E-08	0.002364	2.85E-09	5.48E-07	8.53E-10	1.23E-06	3.18E-06	0.0004482	3.71E-08	0.99718152	1.20E-06	1.51E-08
SCO22404SKAG	1.93E-06	0.0153782	1.26E-06	0.0086689	1.53E-07	6.31E-05	6.54E-05	0.9754591	0.0002516	3.97E-05	7.01E-05	4.95E-07
SCO22434SKAG	1.20E-10	2.41E-06	2.24E-08	5.82E-07	1.45E-11	0.999988	7.38E-08	5.01E-09	2.01E-09	5.41E-08	2.12E-09	8.86E-06
SCO22436SKAG	9.16E-09	2.12E-05	3.05E-07	1.84E-09	1.09E-10	0.0074264	4.04E-08	1.68E-09	2.96E-09	0.99129097	3.67E-09	0.00126111
SCO22437SKAG	8.80E-06	0.0034015	1.96E-07	3.35E-05	2.73E-06	1.61E-06	0.0563978	3.90E-05	0.00332931	0.50092689	0.43582318	3.56E-05
SCO22837SKAG	1.94E-06	0.0001577	3.19E-05	8.47E-08	3.60E-06	0.0276403	3.00E-05	3.74E-06	0.00110724	1.37E-08	7.81E-05	0.97094538
SCO22838SKAG	1.19E-08	0.0008418	6.79E-09	4.43E-08	1.28E-08	0.0397906	0.0002021	1.30E-08	5.69E-07	1.09E-09	4.48E-11	0.95916483
SCO22839SKAG	6.21E-07	0.9271244	6.71E-07	0.0026145	1.44E-06	5.57E-05	0.0008903	0.0682323	0.00102853	4.12E-06	3.67E-08	4.74E-05
SCO22841SKAG	4.02E-07	0.0121598	8.75E-08	1.39E-05	7.48E-06	4.20E-06	0.984774	2.20E-06	0.00195641	1.24E-08	4.00E-05	0.00104162
SCO22845SKAG	1.50E-07	0.1249034	7.55E-07	0.873972	2.50E-08	0.00043	0.000603	3.86E-06	6.88E-05	3.46E-06	1.38E-05	8.00E-07
SCO22846SKAG	1.67E-06	0.015639	0.0011176	7.14E-05	0.0002439	0.0001673	0.9630158	0.0002103	0.01025457	3.98E-06	0.00418215	0.00509243
SCO22847SKAG	9.83E-06	0.3811757	2.29E-07	0.0003786	0.615883	1.68E-05	0.0018152	4.84E-05	0.00045892	7.89E-07	0.00016038	5.21E-05
SCO22848SKAG	2.56E-08	0.001655	8.10E-07	4.54E-05	2.07E-09	0.9752475	3.03E-07	5.19E-08	4.93E-06	4.20E-07	8.81E-10	0.0230456
SCO22849SKAG	8.56E-15	4.53E-10	2.42E-12	9.21E-07	1.10E-15	0.9999983	3.95E-13	2.31E-13	4.13E-11	7.24E-13	4.69E-15	7.32E-07
SCO22850SKAG	4.84E-08	0.000254	3.60E-08	3.41E-06	0.0569589	7.63E-06	0.9401185	6.41E-06	0.00065408	1.40E-08	0.00091286	0.00108408
SCO22852SKAG	5.06E-14	1.28E-09	1.28E-09	8.78E-12	1.46E-11	0.999377	2.20E-08	8.95E-12	3.19E-10	5.19E-14	3.05E-11	0.00062302
SCO22855SKAG	0.0415845	0.0394328	0.571502	4.51E-05	4.08E-06	0.0001945	0.1231205	1.48E-05	0.06752906	0.00010457	0.00057529	0.15589273
SCO22856SKAG	1.21E-14	1.96E-09	7.16E-11	8.75E-13	1.03E-15	0.9999147	4.30E-12	4.44E-14	3.64E-13	1.27E-12	3.14E-15	8.53E-05
SCO22857SKAG	6.18E-09	1.43E-08	6.27E-11	4.49E-09	5.84E-11	2.22E-10	2.05E-06	2.68E-10	7.56E-10	0.99992861	6.93E-05	9.64E-10
SCO22859SKAG	4.02E-06	0.8861236	4.33E-06	0.0212926	6.73E-06	0.0008409	0.0788677	7.90E-05	0.0124275	0.00017748	1.48E-05	0.00016133
SCO22860SKAG	2.29E-08	0.0001276	9.00E-09	2.69E-05	2.75E-08	2.03E-08	2.52E-06	6.62E-08	1.29E-06	0.99984068	7.12E-09	7.96E-07
SCO22861SKAG	1.83E-07	0.3617827	0.0001969	0.6075146	1.41E-06	0.0001382	0.0008008	2.18E-06	0.02559368	1.78E-06	1.18E-08	0.00396757
SCO22862SKAG	1.59E-06	0.0302283	1.93E-05	3.16E-05	3.92E-05	3.13E-05	0.4167765	0.5456355	0.00494855	8.80E-07	0.00066269	0.00162469
SCO22863SKAG	8.10E-14	1.17E-07	0.9998922	1.16E-10	3.20E-10	5.84E-10	3.41E-05	5.21E-10	5.73E-05	2.99E-10	1.56E-05	7.61E-07
SCO22865SKAG	1.06E-07	0.0014532	4.69E-08	2.08E-06	2.51E-08	0.9984862	9.00E-06	9.66E-06	3.22E-07	2.02E-08	7.87E-07	3.86E-05
SCO22866SKAG	1.94E-07	0.0164402	5.68E-05	3.39E-05	1.33E-06	0.0001429	0.9158536	0.00026	0.01360526	0.0125915	0.0409171	9.71E-05
SCO22868SKAG	0.0006008	0.0729928	5.30E-10	0.0024074	1.22E-05	1.53E-06	4.90E-06	0.8945486	0.02941714	5.50E-07	5.36E-06	8.73E-06
SCO22950SKAG	7.48E-09	0.0005488	2.19E-07	0.0251095	1.49E-07	5.05E-06	8.01E-05	2.29E-06	2.38E-05	0.97422985	7.76E-08	1.69E-07
SCO22951SKAG	4.53E-08	0.0012422	1.58E-06	8.86E-08	8.54E-10	0.9985629	1.06E-05	1.36E-06	1.99E-07	4.54E-07	1.94E-06	0.00017857
SCO22952SKAG	1.09E-06	0.0144534	0.0074385	2.50E-06	2.11E-06	5.36E-05	0.0158036	3.53E-06	0.00313013	4.00E-06	2.72E-06	0.9591048
SCO22953SKAG	1.10E-08	0.0010424	1.59E-05	1.33E-07	4.56E-09	0.9611174	2.77E-05	1.16E-07	3.12E-07	5.21E-09	7.80E-09	0.03779604
SCO22954SKAG	7.80E-10	2.69E-06	7.99E-12	7.60E-06	1.17E-09	7.30E-09	2.56E-08	0.9999896	5.69E-08	8.92E-10	4.96E-10	4.78E-09
SCO22955SKAG	7.32E-10	0.0091458	0.9849568	4.94E-06	3.96E-09	7.17E-05	0.0013016	1.68E-06	0.00428195	2.14E-06	0.00020364	2.97E-05
SCO22956SKAG	4.84E-07	0.0062976	0.0001845	0.0018386	7.40E-06	4.60E-06	0.0140364	0.0006931	0.97635548	2.47E-06	0.00056946	9.82E-06
SCO22959SKAG	1.71E-09	0.0012324	0.9910261	0.000116	2.54E-07	7.10E-06	0.0001992	1.58E-07	0.00592013	6.27E-07	4.53E-07	0.00149753
SCO22960SKAG	1.31E-10	1.20E-05	0.0077185	1.05E-07	1.23E-08	0.0332609	0.0053511	1.88E-09	0.00076737	4.00E-09	1.06E-05	0.95287937
19PR0123	9.97E-08	0.0002576	3.29E-06	0.0001296	0.7569613	2.52E-05	0.2384352	2.74E-05	0.00045201	1.06E-07	0.00217899	0.00152912
19PR0134	0.0012243	0.7486195	9.66E-05	0.2350399	0.0004101	0.0006373	0.000183	0.0029731	0.00994423	5.40E-05	0.00027415	0.00054373
19PR0135	5.86E-09	1.02E-06	2.43E-07	0.0017922	2.64E-07	1.27E-09	1.05E-05	0.0113647	0.98485943	1.23E-09	0.00197159	1.26E-08
19PR0138	2.53E-08	0.0007887	1.57E-07	6.22E-06	2.75E-08	5.58E-07	0.0044171	0.0008641	2.99E-05	0.99368575	0.00020085	6.71E-06
19PR0139	1.96E-08	0.0001379	1.28E-07	7.07E-09	0.0006207	0.0242542	1.05E-06	4.51E-09	6.48E-08	5.30E-09	6.04E-10	0.97498598

	1	2	3	4	5	6	7	8	9	10	11	12
19PR0142	2.14E-05	0.0084348	1.91E-07	1.03E-05	0.9900474	5.08E-07	8.35E-06	0.0004445	0.00099735	4.28E-08	2.83E-05	7.00E-06
19PR0152	4.13E-09	0.0002494	1.23E-08	0.9992687	4.17E-10	3.27E-08	2.36E-08	6.04E-08	0.00048173	1.16E-09	4.88E-08	7.56E-10
19PR0155	0.9999988	1.13E-08	1.07E-14	1.19E-06	8.61E-13	4.66E-12	2.56E-12	1.42E-10	6.02E-11	4.57E-12	1.42E-11	2.56E-11
19PR0159	1.25E-09	0.0006149	3.33E-06	3.26E-06	0.0011367	0.996994	0.0001733	4.60E-06	1.92E-05	2.42E-07	2.60E-06	0.00104796
19PR0162	1.03E-05	0.023539	0.0003112	0.0186653	8.29E-05	8.46E-06	0.0172881	0.0002226	0.93792494	9.63E-06	0.00043564	0.00150204
19PR0164	8.63E-07	0.173139	1.50E-06	0.012833	1.21E-06	2.93E-06	0.0004021	7.26E-05	0.81353661	1.24E-06	2.17E-06	6.81E-06
19PR0165	1.35E-08	0.0264497	0.0036236	0.005541	2.19E-07	3.30E-07	0.0008953	5.67E-07	0.96289395	9.27E-08	1.35E-06	0.00059394
19PR0166	2.77E-07	0.0044234	3.15E-05	0.0300298	3.89E-05	5.37E-06	0.0150794	0.801589	0.14822345	1.46E-06	1.14E-05	0.00056601
19PR0172	3.41E-09	1.88E-05	2.26E-05	6.95E-05	1.93E-07	8.44E-08	0.0032381	0.0798199	0.28907265	5.18E-08	0.62775799	1.05E-07
19PR0173	0.2593333	0.0063206	3.68E-08	0.7338271	1.30E-07	2.83E-06	1.40E-06	8.33E-05	0.00029779	4.07E-07	0.00013309	1.98E-08
19PR0174	0.9853025	8.16E-05	5.41E-10	6.43E-06	3.11E-06	1.74E-07	1.78E-07	0.0145979	8.94E-08	2.38E-08	1.11E-06	6.95E-06
19QK0581	4.17E-08	0.0144656	4.21E-08	0.9726285	0.008893	3.55E-06	0.000437	5.78E-06	0.00356158	7.66E-07	1.53E-07	3.98E-06
19QK0590	1.72E-05	0.0027042	0.0001196	0.0002908	8.87E-05	2.52E-07	0.0004242	0.0005232	0.98886833	1.93E-07	0.00042712	0.00653616
19QK0604	2.06E-07	0.0004511	8.02E-09	0.0010498	5.06E-07	2.31E-07	3.78E-07	0.9982468	0.00024495	1.46E-07	5.88E-06	2.00E-09
19QK0629	2.18E-07	0.005279	0.0001611	3.83E-05	2.48E-06	1.13E-05	0.0102373	0.9614596	0.02073056	1.54E-06	0.00186432	0.00021435
19QK0636	3.74E-06	0.001038	8.28E-08	0.0020416	1.82E-06	0.9954018	1.80E-06	1.31E-05	7.10E-06	2.76E-08	1.34E-06	0.00148954
19QK0638	2.11E-09	0.0019829	0.0003434	0.0001614	4.30E-08	2.77E-08	0.0013776	7.06E-08	0.9952439	2.01E-09	0.00089008	6.07E-07
19QK0639	1.80E-09	5.14E-05	4.16E-07	0.9320136	3.81E-09	4.12E-07	1.56E-05	0.0116064	0.05631132	1.25E-07	4.18E-07	3.08E-07
19QK0654	9.22E-07	0.0007894	1.15E-07	0.0935653	0.0001927	1.83E-06	5.98E-05	0.8922809	0.01310549	1.51E-07	2.81E-06	5.83E-07
19QK0761	4.35E-10	0.0002304	0.0016442	0.0004706	7.74E-09	4.86E-08	3.42E-05	5.08E-07	0.99761028	4.11E-08	9.25E-06	5.67E-07
19QK0768	0.0104723	0.0001213	4.28E-11	0.9893992	1.41E-08	2.39E-08	7.76E-09	8.48E-07	5.71E-06	2.03E-09	5.29E-07	1.77E-10
19QK0777	2.80E-09	7.46E-05	7.13E-07	6.21E-06	2.53E-10	3.59E-08	5.53E-08	1.72E-08	1.98E-06	0.99991634	1.22E-08	2.08E-08
19QK1175	2.06E-08	0.0037111	0.1366322	0.004287	6.75E-08	1.13E-05	0.000103	4.92E-06	0.8551452	4.19E-06	8.76E-05	1.34E-05
19QK1182	0.0009778	0.7133236	0.005753	0.2098199	0.0003966	0.0095691	0.0025829	0.016403	0.03250807	0.0004305	0.00199942	0.00623616
19QK1183	2.72E-06	0.0004638	3.42E-08	0.0011914	3.64E-06	3.89E-08	2.72E-05	0.4028832	0.59525966	2.72E-08	0.00016297	5.28E-06
19QK1188	1.02E-06	0.0218919	0.0034936	0.0026875	3.07E-05	3.87E-05	0.0488831	2.38E-05	0.07539251	2.90E-07	2.13E-06	0.84755468
19QK1236	3.78E-08	9.31E-05	2.23E-08	0.9050702	2.11E-08	1.22E-07	1.10E-07	7.67E-07	0.00039531	0.09443988	3.47E-07	2.22E-08
19QK1254	1.02E-10	0.0014186	5.41E-05	0.0378653	0.004349	1.31E-07	0.0290985	9.05E-07	0.92719011	2.68E-08	1.85E-05	4.95E-06
19QK1395	1.08E-07	0.0139632	0.0001977	6.05E-05	8.19E-06	2.07E-07	0.0015859	0.0002463	0.98372378	9.29E-08	0.00012946	8.45E-05
19QK1415	0.0064081	0.0123044	0.0008845	0.0013093	0.7119044	0.000106	0.1048439	0.0024942	0.12429499	8.58E-06	0.02634309	0.00909866
19QK1422	0.0075806	0.0002871	0.000298	6.89E-06	1.48E-07	0.3712674	0.0007672	0.0003097	0.00230418	1.60E-06	0.61619887	0.00097813
19QK1443	4.32E-09	6.53E-06	2.77E-07	5.02E-09	0.9984398	2.76E-10	4.79E-05	1.60E-06	0.0014464	5.38E-11	5.69E-05	4.56E-07
SCO22120SKAG	5.16E-09	0.0001509	3.63E-11	0.9995633	5.65E-10	1.03E-06	1.55E-06	5.33E-08	0.0002828	6.07E-09	5.02E-10	2.99E-07
SCO22122SKAG	7.70E-07	0.0092927	8.20E-07	0.2562193	2.02E-06	0.0002062	0.5797546	3.05E-06	0.13270095	8.73E-07	2.99E-05	0.02178881
SCO22124SKAG	3.56E-09	0.0001037	1.07E-10	0.9994994	0.0003682	2.96E-07	1.44E-06	7.84E-08	2.69E-05	7.78E-09	8.15E-10	3.32E-08
SCO22126SKAG	3.48E-05	5.24E-05	2.92E-08	5.94E-08	0.0044279	1.67E-09	0.9776765	1.95E-09	0.00333085	2.17E-10	0.00035755	0.01411989
SCO22142SKAG	3.66E-06	0.0064996	2.70E-07	0.0327183	0.000139	0.000147	0.5490394	0.2545736	0.13524457	1.16E-06	1.21E-05	0.02162136
SCO22143SKAG	4.62E-08	0.0001016	1.50E-06	2.08E-07	0.9816997	6.89E-09	0.000636	9.18E-06	0.01296089	6.35E-10	0.00458648	4.32E-06
SCO22145SKAG	1.82E-06	0.0014263	5.04E-08	0.2220076	2.49E-07	4.83E-06	1.22E-05	3.04E-06	7.83E-05	0.77645842	4.16E-07	6.76E-06
SCO22146SKAG	4.56E-08	0.0006346	1.72E-08	0.9989729	3.37E-08	5.54E-07	2.76E-07	1.22E-06	0.00039029	2.96E-08	3.11E-09	1.13E-08
SCO22148SKAG	3.88E-07	0.0150878	2.46E-07	0.0536085	2.14E-07	1.90E-06	0.0001746	0.0001863	0.93093286	1.64E-07	3.93E-07	6.70E-06
SCO22149SKAG	2.47E-05	0.5976489	0.002834	0.000293	0.0002644	9.10E-06	0.0006305	0.0007384	0.39552877	5.98E-06	5.06E-06	0.00201719
SCO22729SKAG	3.48E-09	0.0037411	4.50E-09	0.0136369	0.0454091	1.36E-06	0.8790729	0.0128169	0.04529238	1.17E-07	2.17E-05	7.57E-06
SCO22731SKAG	1.27E-05	0.3183672	1.16E-07	1.37E-05	4.96E-05	0.000522	0.1541156	2.05E-06	5.77E-05	1.14E-06	1.92E-07	0.52685803
SCO22737SKAG	5.82E-10	0.0696554	2.91E-06	0.0001427	2.83E-08	1.23E-05	0.9289393	1.14E-07	0.00123423	2.06E-08	2.60E-06	1.04E-05
SCO22743SKAG	3.67E-11	3.26E-07	2.21E-10	4.64E-08	5.46E-13	6.12E-09	4.74E-09	3.10E-10	2.67E-08	0.99999959	3.08E-10	1.97E-09
SCO22754SKAG	4.76E-08	0.0179747	3.48E-06	2.42E-07	1.45E-07	1.44E-06	0.9588546	1.52E-06	0.02161796	4.46E-07	5.50E-05	0.0014904
SCO22772SKAG	0.0002801	0.9607195	1.03E-07	0.0028425	1.51E-05	0.0001495	0.0116362	0.0001215	0.02285898	1.47E-06	3.25E-06	0.00137188
SCO22786SKAG	2.53E-08	0.0132424	3.14E-07	2.01E-05	5.76E-08	0.3347368	0.1059934	5.00E-08	0.00154594	1.42E-08	1.90E-05	0.54444196
SCO22790SKAG	2.16E-07	0.0053896	2.04E-08	3.27E-08	6.21E-09	0.0874601	3.87E-05	1.23E-08	1.51E-06	1.96E-08	3.97E-10	0.90710979

	1	2	3	4	5	6	7	8	9	10	11	12
SCO22791SKAG	7.18E-07	0.1033268	3.22E-06	5.39E-07	1.18E-06	1.42E-05	0.1828838	4.55E-08	0.0003912	1.30E-06	1.90E-07	0.71337671
SCO22792SKAG	2.76E-06	0.8162939	0.0306611	2.52E-06	1.50E-05	5.40E-05	0.0973877	5.71E-05	0.04073556	8.28E-06	0.00777694	0.00700509
SCO22793SKAG	3.20E-09	2.51E-05	3.12E-05	6.43E-05	8.25E-07	7.52E-06	0.9441899	3.21E-08	0.00123862	1.57E-09	3.64E-05	0.05440601
SCO22794SKAG	2.28E-07	0.0304365	5.72E-07	3.73E-06	3.68E-06	4.73E-06	0.9630789	5.22E-06	0.00604148	1.78E-08	6.11E-05	0.00036381
SCO22795SKAG	5.97E-10	2.42E-06	1.35E-09	5.55E-08	6.66E-12	4.60E-09	2.90E-08	4.76E-09	8.54E-08	0.99999739	1.98E-09	2.94E-09
SCO22796SKAG	1.05E-06	0.9275401	5.82E-05	0.0014789	7.78E-07	0.0019701	0.0631804	0.0029366	0.00266438	1.39E-05	3.47E-07	0.00015522
SCO22798SKAG	1.34E-07	0.0079064	2.47E-05	1.06E-06	2.50E-08	0.9168839	1.27E-05	9.79E-07	1.14E-05	3.67E-07	5.59E-09	0.07515832
SCO22799SKAG	3.27E-10	1.24E-05	1.60E-08	2.06E-07	1.80E-11	3.00E-09	6.77E-09	2.35E-09	1.43E-07	0.99998721	2.52E-09	1.25E-09
SCO22800SKAG	6.65E-11	0.0001228	2.75E-07	1.51E-07	0.0121783	6.29E-07	0.9870668	1.11E-06	0.00048298	2.74E-08	0.00014534	1.59E-06
SCO22801SKAG	4.15E-06	0.0015112	1.45E-07	0.0048673	1.62E-07	4.14E-07	0.0001221	7.88E-05	0.00677972	0.30452121	0.68211483	2.75E-08
SCO22802SKAG	5.19E-10	0.0001082	5.87E-07	0.001311	1.38E-09	0.9976745	0.0006423	8.53E-08	2.42E-05	7.69E-09	8.58E-06	0.00023041
SCO22803SKAG	3.29E-06	0.6894304	6.97E-05	5.24E-05	5.31E-05	0.0003017	0.2075566	0.0933801	0.00246812	8.55E-06	1.69E-05	0.00665907
SCO22807SKAG	2.00E-09	0.002959	4.10E-09	8.64E-06	0.0085352	1.76E-06	0.9878161	3.91E-07	0.00066073	1.26E-08	1.37E-05	4.38E-06
SCO22808SKAG	5.94E-07	0.9991374	1.32E-06	6.76E-05	3.34E-07	1.54E-05	0.0006622	1.50E-06	8.15E-05	1.01E-06	3.44E-07	3.07E-05
19PR0111	1.68E-05	0.0216698	8.18E-06	0.9731093	5.22E-05	3.55E-05	8.06E-06	4.74E-05	0.00498538	3.47E-07	1.27E-07	6.70E-05
19PR0114	1.40E-10	3.05E-05	2.00E-09	0.9978807	8.03E-05	3.91E-08	5.66E-05	1.34E-08	0.00195158	5.71E-10	2.14E-07	3.28E-08
19PR0115	3.83E-08	2.86E-05	8.08E-08	2.54E-09	4.26E-07	0.0030329	0.0003696	2.74E-08	6.36E-05	4.72E-10	2.96E-07	0.99650434
19PR0116	7.26E-07	0.2556637	0.3188082	0.0001387	0.0001793	7.49E-05	0.1667966	8.64E-05	0.24908579	2.27E-05	0.00024231	0.00890072
19PR0119	4.67E-11	8.96E-05	0.9996115	8.84E-09	1.19E-09	6.36E-08	8.80E-06	6.01E-09	0.00023577	1.52E-07	4.83E-08	5.41E-05
19PR0120	5.92E-11	3.88E-05	3.34E-05	0.0004757	5.35E-08	3.90E-07	0.9021029	4.30E-07	0.05324358	0.03543457	0.00866698	3.18E-06
19PR0149	1.61E-10	4.60E-05	3.74E-07	0.9641134	2.71E-09	2.80E-08	8.24E-06	9.04E-08	0.03583174	1.67E-09	1.09E-07	2.08E-08
19PR0176	1.09E-10	0.0033807	0.9903294	5.50E-07	3.26E-08	4.23E-07	0.0017861	9.82E-08	0.00397614	1.06E-08	7.14E-08	0.00052651
19PR0177	2.44E-08	0.0367789	0.0002562	0.00103	2.77E-06	9.60E-05	0.9304472	5.56E-06	0.0287123	1.48E-07	4.58E-06	0.00266634
19PR0178	7.76E-07	0.9476868	0.0008066	0.0188537	9.94E-07	0.0026853	0.0055536	4.17E-06	0.00337587	6.23E-06	2.65E-08	0.02102593
19PR0179	2.57E-09	0.147921	0.0039288	8.24E-05	5.08E-07	8.62E-07	0.0373871	2.17E-06	0.81064103	3.99E-08	2.21E-05	1.39E-05
19PR0180	1.47E-07	0.0141748	6.16E-05	0.0009018	0.465951	1.73E-06	0.0083371	4.21E-05	0.50632238	4.52E-07	0.00418595	2.09E-05
19PR0181	1.05E-05	0.0623607	4.05E-05	3.79E-06	0.8832886	2.22E-06	6.19E-05	1.22E-06	0.00079776	1.37E-06	8.61E-07	0.05343063
19PR0182	1.25E-06	0.0475778	1.31E-05	0.0005908	0.4080432	0.0001244	0.0014481	6.08E-07	0.0003565	1.20E-06	4.82E-08	0.54184298
19QK1284	1.71E-09	5.19E-05	1.42E-10	0.9995683	1.12E-10	8.51E-08	7.38E-08	0.0003609	1.88E-05	8.86E-09	5.11E-09	3.26E-10
19QK1293	1.85E-10	9.15E-05	0.0004757	0.0003796	4.96E-08	1.29E-08	0.3418973	6.38E-08	0.6378194	0.01621782	0.00295625	0.00016234
SCO22247SKAG	4.64E-09	0.0006978	0.5518533	6.99E-07	3.34E-06	4.61E-07	0.0436585	9.90E-09	0.09687632	1.34E-08	6.84E-05	0.30684108
SCO22249SKAG	1.06E-07	6.10E-05	5.05E-10	0.0001716	9.32E-10	2.50E-06	3.07E-06	9.01E-09	6.40E-07	0.99974324	3.04E-08	1.78E-05
SCO22252SKAG	6.68E-10	0.0002858	2.01E-07	2.59E-07	2.84E-07	1.97E-07	0.935243	2.71E-07	0.00067142	0.06375114	4.20E-05	5.42E-06
SCO22254SKAG	9.63E-06	4.90E-07	4.85E-10	0.9992361	1.12E-10	4.85E-10	4.63E-08	3.45E-08	0.00075356	1.93E-10	1.70E-07	5.51E-10
SCO22255SKAG	0.9825778	0.0008858	1.23E-08	1.16E-06	2.98E-07	0.0132863	3.23E-06	2.57E-05	0.00010032	1.42E-07	2.79E-05	0.00309133
SCO22256SKAG	0.001059	0.0002642	5.36E-09	6.90E-06	5.92E-10	1.88E-07	3.45E-06	1.23E-07	3.48E-08	0.9986658	2.44E-07	5.62E-08
SCO22010SKAG	5.85E-07	0.0243252	1.36E-07	0.0001955	1.38E-08	0.9278656	6.12E-05	9.44E-08	2.31E-05	3.62E-07	2.63E-07	0.0475279
SCO22013SKAG	1.75E-13	4.00E-09	1.75E-13	3.80E-12	0.9999957	1.15E-12	3.67E-06	3.00E-13	1.90E-09	6.39E-15	4.68E-11	6.42E-07
SCO22016SKAG	6.12E-06	0.9912881	0.0001177	0.0033742	1.72E-06	1.68E-05	1.68E-05	3.49E-06	0.00404615	7.73E-06	2.81E-08	0.0011211
20NW0260	1.76E-07	3.71E-05	5.06E-08	5.15E-09	0.9973063	6.62E-09	0.0001483	0.0021252	3.97E-05	1.67E-09	4.46E-06	0.00033871
20NW0261	0.0098432	8.79E-07	9.49E-08	1.36E-07	2.47E-07	8.93E-10	7.58E-05	0.0384088	0.00096113	7.50E-10	0.95070906	6.22E-07
20NW0262	2.39E-08	0.0002117	5.59E-07	3.87E-08	0.0013991	1.31E-07	0.0008304	1.33E-06	4.05E-05	0.68823164	0.30928434	2.18E-07
20NW0264	4.75E-06	0.0013449	1.25E-09	1.40E-05	0.9877921	8.47E-07	9.16E-06	0.0104906	8.35E-06	2.95E-08	1.32E-07	0.00033511
20NW0268	0.2901679	0.0454669	4.06E-07	0.0004029	0.6616173	2.21E-05	0.0010451	0.0005885	0.000222	1.85E-05	0.00044288	5.52E-06
20NW0269	0.9721665	0.0001437	4.67E-11	4.39E-07	1.37E-06	3.62E-09	1.19E-08	0.0276646	3.51E-07	7.13E-09	2.30E-05	7.10E-09
20NW0272	7.44E-07	0.0004459	3.65E-07	7.90E-07	2.41E-06	0.1584259	5.77E-07	0.8410588	3.89E-05	3.67E-07	5.32E-06	2.00E-05
20NW0275	2.03E-06	0.0013634	1.59E-05	2.54E-05	5.93E-05	1.36E-05	0.0020307	0.9788942	0.00506605	4.90E-07	0.01244332	8.58E-05
20NW0279	0.5947097	0.0210257	6.30E-07	3.37E-05	7.31E-07	0.378338	3.32E-06	3.78E-05	2.14E-05	4.65E-07	6.45E-06	0.005822
20NW0281	6.51E-07	0.0017337	0.0004326	4.49E-05	9.65E-05	3.70E-08	0.0002189	4.10E-05	0.99716591	6.21E-09	0.00019011	7.58E-05
20NW0284	3.35E-06	0.0237577	4.51E-08	0.0009748	0.9736883	2.23E-07	1.17E-05	1.79E-05	0.00152169	8.03E-08	4.89E-07	2.38E-05

	1	2	3	4	5	6	7	8	9	10	11	12
20NW0287	0.000395	0.0141295	3.65E-07	0.0001256	0.0003248	5.84E-06	7.82E-05	0.0003178	2.63E-05	0.98362166	0.00034006	0.00063483
20NW0288	0.4400432	0.0078908	0.0081801	0.0001317	5.18E-06	2.16E-05	0.0137075	0.0180504	0.27985295	0.00034817	0.23171482	5.37E-05
20NW0297	5.82E-07	0.0002752	9.27E-09	1.19E-06	4.46E-07	7.77E-07	4.02E-06	0.9997098	4.98E-06	7.61E-08	2.93E-06	1.17E-08
20NW0298	0.1275051	0.0003216	4.04E-06	4.07E-06	1.77E-05	4.74E-07	0.0005728	0.79908	0.00050824	9.72E-07	0.07197852	6.50E-06
20NW0301	0.9999992	5.77E-10	3.63E-13	8.79E-12	1.09E-11	3.53E-14	4.24E-11	1.54E-10	1.21E-10	1.12E-13	8.28E-07	2.01E-10
20NW0302	3.89E-08	3.73E-06	1.13E-08	3.37E-09	0.9970875	1.69E-09	4.49E-05	6.67E-06	3.26E-05	8.66E-11	0.00282444	8.11E-08
20NW0307	0.9999999	9.86E-08	2.04E-12	4.35E-08	8.03E-12	3.90E-12	9.90E-12	3.38E-09	1.13E-09	1.15E-10	6.97E-10	1.98E-12
20NW0308	0.990717	9.33E-05	5.43E-09	5.97E-07	9.98E-08	0.0052341	3.22E-08	1.00E-05	9.86E-08	7.34E-09	0.00390608	3.86E-05
20NW0311	1.29E-06	0.0055943	5.61E-07	0.8692179	3.22E-06	0.0001684	1.99E-05	0.1227988	0.00156579	4.90E-06	1.65E-08	0.00062497
20NW0312	9.10E-11	1.14E-08	5.52E-15	6.82E-11	1	1.38E-12	5.66E-10	4.53E-10	7.43E-10	7.96E-14	4.58E-10	2.97E-10
20NW0313	0.999999	2.96E-07	3.93E-12	6.88E-07	1.36E-09	1.22E-09	7.60E-09	3.06E-09	2.10E-10	2.46E-10	3.51E-10	1.40E-10
20NW0314	5.42E-06	0.0001748	2.25E-08	6.99E-08	5.66E-07	0.0134399	6.97E-08	0.9862954	4.28E-06	1.44E-08	4.43E-05	3.53E-05
20NW0321	0.9897365	2.58E-09	8.26E-11	4.78E-11	9.56E-09	9.31E-14	2.57E-07	1.90E-09	1.57E-07	4.72E-13	0.01026309	2.40E-10
20NW0322	0.4057012	9.74E-05	4.73E-09	1.52E-07	5.38E-08	4.62E-09	1.81E-08	1.08E-05	9.83E-07	0.59410294	8.64E-05	3.24E-08
20NW0326	2.10E-07	0.0001796	3.52E-10	6.27E-06	0.1063506	2.25E-06	0.0004383	0.8926629	0.00035388	9.86E-08	5.11E-06	8.26E-07
20NW0327	1.40E-09	5.38E-05	0.999941	2.29E-09	2.19E-10	1.70E-07	4.96E-09	2.62E-06	4.88E-07	8.71E-09	1.72E-06	2.58E-07
20NW0328	7.13E-09	0.0038459	2.12E-06	1.51E-06	8.34E-09	0.9960217	2.24E-05	5.06E-06	1.14E-05	1.17E-07	3.37E-08	8.97E-05
20NW0330	3.38E-11	2.40E-06	0.8895338	1.35E-09	6.63E-09	1.52E-10	7.48E-06	7.59E-08	0.08993604	4.92E-11	0.02051974	4.08E-07
20NW0333	3.59E-10	6.30E-08	1.13E-07	7.80E-09	0.0013549	9.83E-11	0.0005982	3.14E-07	0.00247948	0.00049872	0.99506819	1.42E-08
20NW0334	2.28E-06	8.53E-05	1.28E-09	2.98E-07	3.34E-06	9.48E-08	4.59E-07	0.9985673	3.19E-06	1.17E-08	0.00133776	4.89E-10
20NW0343	0.9999998	4.97E-09	2.32E-14	5.07E-11	1.75E-12	2.26E-13	1.85E-08	3.76E-11	6.29E-10	3.53E-13	1.88E-07	4.12E-09
20NW0345	0.9605891	0.0001187	7.63E-10	1.07E-05	6.14E-07	2.25E-07	2.09E-07	0.0392412	3.29E-07	2.38E-08	2.32E-08	3.88E-05
20NW0346	0.1308303	1.38E-05	1.31E-10	6.73E-07	4.20E-08	9.88E-09	1.24E-08	0.8690758	6.96E-07	1.16E-08	7.87E-05	2.78E-10
20NW0347	5.26E-09	8.66E-08	3.00E-09	4.19E-11	0.0563658	7.13E-12	0.0002595	4.22E-08	1.41E-05	6.69E-06	0.94335296	8.21E-07
20NW0348	0.0002236	0.004641	1.12E-07	0.0037145	3.22E-05	5.07E-06	2.71E-07	0.9899914	0.0013793	6.34E-06	2.10E-06	4.05E-06
20NW0349	5.11E-06	0.0046464	4.67E-05	0.0002457	1.52E-05	3.86E-06	0.0006583	0.8074716	0.18174806	7.44E-06	0.00514425	7.40E-06
20NW0352	0.9670612	0.0024989	3.75E-08	0.0275503	9.53E-08	4.46E-07	3.52E-07	0.0004673	0.00239124	1.18E-06	2.89E-05	2.67E-08
20NW0353	1.72E-11	8.40E-09	8.65E-09	8.65E-14	7.32E-11	3.41E-10	4.63E-06	5.66E-10	7.79E-08	0.00557289	0.99442238	1.01E-08
20NW0354	1.33E-07	4.85E-05	3.38E-10	1.04E-07	0.0159229	1.79E-08	4.24E-07	0.9839469	7.23E-05	1.24E-08	8.75E-06	2.15E-09
20NW0355	3.14E-08	1.44E-05	1.51E-10	2.70E-09	3.27E-07	5.38E-09	1.04E-08	0.9999852	3.60E-08	4.98E-09	2.13E-10	1.61E-08
20NW0357	2.42E-06	0.0002704	1.84E-08	4.37E-07	0.8296618	0.038778	1.22E-05	0.0004732	5.01E-06	1.02E-08	0.13078726	9.22E-06
20NW0373	1.72E-09	0.000477	2.95E-08	1.57E-07	6.45E-08	3.11E-07	0.9992006	2.39E-07	0.00029445	1.27E-09	4.53E-06	2.26E-05
20NW0374	0.0515002	0.0021915	1.06E-07	1.68E-05	0.0172481	3.99E-07	0.0009623	5.98E-05	0.0030143	7.59E-07	0.92500542	3.96E-07
20NW0378	1.01E-08	1.07E-05	1.29E-10	2.44E-07	0.9934517	3.42E-09	0.0061721	1.82E-07	0.00034557	2.76E-11	1.80E-06	1.77E-05
20NW0383	6.63E-06	0.0571966	0.0099624	2.24E-06	3.71E-05	3.20E-06	0.018024	0.0001474	0.13864163	3.19E-07	0.77532618	0.00065221
20NW0386	0.9980598	0.0002455	1.62E-09	2.04E-07	5.49E-08	1.42E-08	2.21E-08	7.76E-05	5.16E-07	5.25E-09	0.00161633	9.92E-09
20NW0391	0.9950756	0.0047362	1.04E-08	1.13E-06	1.22E-05	5.62E-07	9.01E-05	3.07E-05	4.05E-05	3.50E-07	1.52E-06	1.12E-05
20NW0409	0.9953669	0.0029618	1.02E-09	8.69E-08	1.79E-07	1.15E-08	3.31E-08	0.001667	1.87E-07	3.30E-08	1.50E-06	2.30E-06
20NW0415	0.9047471	0.0003326	3.13E-11	1.20E-07	1.96E-06	6.41E-09	3.82E-07	0.0948874	2.71E-05	3.03E-08	3.18E-06	1.07E-07
20NW0432	8.97E-05	0.6652749	6.65E-06	0.3164835	1.89E-05	0.0004323	0.0023603	0.000793	0.01334385	9.40E-06	9.06E-07	0.00118672
20NW0433	1.10E-05	0.0639261	5.16E-06	2.32E-06	6.07E-07	3.90E-05	0.0141039	5.83E-08	0.00040532	6.81E-06	3.10E-07	0.92149935
20NW0434	4.75E-06	0.0002592	6.77E-08	2.95E-07	0.9756818	4.80E-08	6.43E-05	0.0213668	0.00167515	5.37E-08	0.00058426	0.00036334
20NW0436	1.70E-06	0.001646	8.94E-08	6.65E-06	3.36E-06	2.08E-07	4.62E-06	0.9883059	0.01002657	7.35E-08	4.68E-06	2.13E-07
20NW0437	1.77E-11	4.41E-10	2.11E-16	4.10E-13	1	4.18E-14	4.14E-11	2.29E-11	4.57E-12	1.76E-15	4.53E-10	7.93E-11
20NW0438	5.54E-10	8.44E-06	8.94E-07	1.03E-08	1.59E-09	9.04E-08	9.31E-05	0.00049	1.31E-06	0.99865755	0.00074851	3.14E-08
20NW0439	9.87E-11	1.88E-08	1.30E-11	2.32E-11	1.78E-12	0.9999834	1.31E-10	4.80E-10	5.93E-12	3.83E-13	8.16E-09	1.66E-05
20NW0441	1	1.18E-09	2.78E-12	7.14E-12	3.52E-11	6.75E-14	2.94E-11	3.22E-11	4.50E-10	2.55E-12	2.50E-08	9.09E-10
20NW0443	5.45E-06	0.1130539	0.0066061	0.0006187	0.0001653	0.0007516	0.2895432	0.0054841	0.1750719	1.95E-05	0.40861083	6.94E-05
20NW0444	1.99E-08	0.0004395	7.31E-09	0.0010238	3.88E-09	8.19E-07	9.41E-06	4.24E-07	5.22E-07	0.99852436	1.14E-06	1.20E-08
20NW0445	4.79E-08	5.91E-11	2.15E-11	1.29E-12	1.47E-05	7.59E-14	1.19E-06	4.35E-08	2.06E-07	3.04E-14	0.99998385	5.56E-13

	1	2	3	4	5	6	7	8	9	10	11	12
20NW0446	9.86E-06	0.9895542	6.40E-05	0.0005901	7.52E-05	0.000378	0.0052591	0.0001307	0.00166847	5.43E-06	1.27E-07	0.00226474
20NW0447	2.03E-10	1.47E-05	0.00026	4.52E-09	1.45E-10	1.52E-07	9.55E-08	3.29E-08	2.81E-08	0.99972441	5.54E-07	9.18E-09
20NW0448	0.0175903	0.0224039	0.5579832	4.55E-05	2.04E-07	2.76E-05	0.0023408	0.0002211	0.01552854	3.79E-05	0.38380587	1.50E-05
20NW0449	2.04E-11	7.78E-09	3.18E-08	6.93E-11	7.30E-08	1.27E-10	1.84E-05	0.9998971	7.41E-05	5.62E-11	1.03E-05	6.73E-08
20NW0450	1.06E-07	0.0035292	4.32E-06	1.63E-05	0.9948085	1.86E-06	0.0001672	0.0002369	0.00122233	4.49E-07	1.22E-05	5.78E-07
20NW0452	1.23E-06	0.0003604	1.08E-05	5.18E-07	0.0002594	1.31E-06	0.0017087	0.9918452	0.00087679	1.95E-07	0.00492907	6.41E-06
SCO22023SKAG	4.69E-06	0.0009231	1.13E-06	7.65E-07	0.0003124	7.47E-08	9.17E-07	0.9976844	0.00107008	5.96E-08	1.40E-06	8.64E-07
SCO22025SKAG	2.38E-06	0.0003106	3.26E-08	4.80E-06	1.12E-06	5.49E-08	5.37E-08	0.9995022	0.00015854	2.64E-08	2.01E-05	5.14E-08
SCO22026SKAG	0.0016769	4.98E-06	7.63E-12	0.9954726	6.91E-10	1.41E-08	8.98E-10	0.0028403	5.23E-06	5.48E-09	3.67E-09	2.47E-10
SCO22027SKAG	1.57E-06	0.1374216	2.78E-05	0.0172011	0.8062867	0.000298	0.0050555	0.000516	0.03301263	2.89E-06	4.79E-05	0.00012833
SCO22028SKAG	4.42E-05	0.0017989	5.56E-07	2.10E-07	1.04E-05	1.37E-07	5.00E-07	0.0001265	5.11E-05	0.99790519	6.13E-05	1.03E-06
SCO22029SKAG	5.70E-10	2.67E-08	5.45E-15	5.94E-11	7.41E-10	7.54E-12	5.11E-11	1	6.45E-09	6.03E-12	7.98E-09	1.44E-13
SCO22030SKAG	0.9940438	0.0002745	3.01E-09	2.79E-08	1.70E-05	1.37E-08	1.30E-07	0.0056589	5.85E-08	3.43E-08	1.17E-06	4.30E-06
SCO22031SKAG	0.9989414	0.0010477	1.76E-09	1.63E-06	6.44E-07	1.52E-08	2.50E-08	5.13E-06	2.83E-06	1.24E-08	4.17E-07	1.44E-07
SCO22032SKAG	0.089717	0.0001342	3.46E-06	4.15E-05	1.48E-06	1.08E-07	1.30E-05	0.8748477	0.01690761	6.69E-07	0.01833301	1.76E-07
SCO22033SKAG	0.999804	0.0001914	2.86E-09	1.47E-07	6.42E-08	2.57E-09	3.69E-09	6.49E-07	1.05E-06	2.74E-08	1.18E-07	2.55E-06
SCO22034SKAG	8.39E-05	0.0043185	0.0005178	2.38E-06	0.0008669	0.2506738	0.0026343	0.0005594	0.01170693	7.96E-07	0.54404281	0.18459267
SCO22035SKAG	0.6718556	0.3280746	1.28E-08	2.22E-05	2.32E-08	1.04E-06	5.93E-06	7.34E-06	7.16E-06	3.16E-07	2.57E-05	1.09E-07
SCO22036SKAG	4.10E-12	6.91E-10	2.71E-09	6.43E-12	4.60E-11	5.49E-07	1.32E-06	9.72E-10	6.03E-08	1.01E-06	0.99999704	9.42E-09
SCO22037SKAG	1.92E-09	6.65E-07	1.47E-11	1.39E-09	2.94E-09	1.84E-09	7.55E-10	0.9999993	1.41E-08	8.90E-10	9.55E-09	7.59E-10
SCO22038SKAG	1.89E-07	3.77E-05	3.14E-07	6.35E-08	2.31E-08	1.07E-08	2.84E-08	3.56E-05	2.66E-05	0.99985871	4.08E-05	1.08E-09
SCO22040SKAG	0.1932412	0.0205868	0.0002453	0.0805664	1.35E-05	8.30E-06	0.0106981	0.0016189	0.67476594	1.15E-05	0.01823493	9.15E-06
SCO22044SKAG	2.22E-05	0.9981564	3.30E-07	4.33E-05	9.83E-06	2.20E-05	0.0004897	1.70E-05	0.00098512	5.35E-07	1.12E-07	0.00025351
SCO22047SKAG	0.9999996	4.22E-08	1.12E-12	4.24E-08	5.91E-12	3.35E-11	7.54E-08	1.86E-10	8.08E-08	2.29E-10	1.66E-07	5.04E-09
SCO22048SKAG	0.8597767	0.1373761	3.91E-07	0.0015204	8.14E-08	3.19E-05	4.60E-05	0.0003057	9.03E-05	1.06E-05	0.0008417	1.65E-07
SCO22049SKAG	0.4934088	0.1197388	6.06E-05	4.65E-05	1.09E-05	2.95E-06	0.0165673	0.0004566	0.02167364	4.67E-07	0.34794518	8.82E-05
SCO22052SKAG	0.9996206	0.0002053	7.23E-09	0.0001487	2.35E-07	2.26E-08	8.03E-09	4.46E-06	1.82E-05	1.35E-06	7.99E-07	2.29E-07
SCO22053SKAG	0.0025048	0.5182122	1.61E-07	0.1513336	0.0013608	1.05E-05	0.0001398	0.0227226	0.30224503	8.21E-07	0.00142157	4.82E-05
SCO22054SKAG	0.0014109	0.5328082	0.0022068	0.0018089	0.0001061	0.0001491	0.0002781	0.0600141	0.34737966	0.00034191	0.05345829	3.79E-05
SCO22055SKAG	6.77E-12	1.31E-08	1.76E-14	3.59E-11	0.9999996	1.03E-12	2.73E-07	1.91E-11	1.96E-09	3.63E-15	1.11E-07	1.01E-08
SCO22059SKAG	0.982369	0.0007374	2.15E-08	8.47E-06	7.42E-06	6.27E-08	0.0009145	0.0001104	0.01357263	2.04E-08	0.00227132	8.72E-06
SCO22060SKAG	0.4743804	0.0053295	3.57E-09	0.0001157	0.5182312	8.14E-07	0.000102	2.66E-05	6.22E-06	3.64E-07	0.00180681	4.62E-07
SCO22062SKAG	0.9999997	2.15E-07	2.20E-13	4.25E-08	6.68E-12	3.92E-11	9.67E-11	8.44E-10	1.98E-10	6.00E-11	1.24E-08	3.04E-12
SCO22065SKAG	0.9899821	0.0001016	3.72E-12	2.97E-08	2.07E-07	5.61E-10	1.85E-07	5.14E-07	5.17E-07	0.00990439	1.04E-05	5.45E-08
SCO22068SKAG	1.16E-11	1.73E-10	1.17E-14	2.49E-14	1.91E-13	7.58E-15	7.27E-14	1.89E-11	1.50E-12	1	9.27E-10	3.01E-14
SCO22069SKAG	7.22E-07	0.0358243	0.0060418	6.19E-06	1.72E-07	0.9168188	4.01E-07	1.18E-06	4.79E-05	2.16E-07	8.83E-08	0.04125819
SCO22072SKAG	0.0095747	0.9153586	0.0001074	0.0001132	0.0021798	0.0001126	4.41E-05	0.0020765	0.00942267	1.62E-05	3.53E-05	0.06095912
SCO22073SKAG	1.09E-08	0.00016	0.0004819	2.31E-08	1.00E-08	6.36E-08	3.95E-08	7.49E-07	2.18E-06	0.999354	9.67E-07	1.32E-08
SCO22074SKAG	2.50E-05	0.0097572	0.0001936	4.65E-07	3.54E-07	1.96E-07	6.34E-08	1.53E-05	1.96E-05	0.98856604	0.00142136	8.42E-07
SCO22075SKAG	0.9957542	0.003583	0.0005725	1.48E-06	1.88E-05	1.47E-07	6.79E-08	2.14E-05	3.77E-05	6.42E-07	4.35E-06	5.67E-06
SCO22076SKAG	7.07E-05	0.0133828	7.41E-07	3.10E-05	2.66E-05	1.09E-05	1.07E-05	0.9862247	0.0001444	2.45E-06	2.12E-07	9.49E-05
SCO22077SKAG	0.022166	0.0418253	0.9351249	1.16E-06	5.49E-07	1.24E-05	4.76E-06	3.28E-06	3.79E-05	0.00012771	5.18E-07	0.00069547
SCO22079SKAG	3.59E-06	0.002108	8.30E-10	5.10E-07	0.9904874	1.95E-07	0.0002005	0.0012902	0.00144217	2.10E-08	0.00446742	4.60E-08
SCO22080SKAG	3.81E-08	0.000282	1.49E-05	6.88E-05	0.9747096	2.13E-07	0.0041906	1.01E-06	0.02035738	5.54E-09	2.05E-05	0.00035491
SCO22081SKAG	2.63E-06	0.0002458	3.46E-09	1.10E-05	0.8567379	5.90E-07	3.60E-05	5.18E-06	1.43E-06	0.14293016	7.79E-06	2.15E-05
SCO22082SKAG	0.9948048	0.0030083	0.0004803	0.0011659	7.56E-07	3.51E-07	1.14E-08	4.40E-05	0.00048537	1.95E-06	4.15E-06	4.19E-06
SCO22083SKAG	3.20E-07	0.0002012	0.0011937	1.94E-08	5.05E-08	4.48E-08	6.72E-09	4.71E-06	1.38E-05	0.99857503	1.07E-05	4.43E-07
SCO22084SKAG	1	3.96E-08	6.36E-12	3.17E-11	1.09E-11	1.51E-12	5.08E-12	1.98E-10	4.65E-11	4.68E-11	1.47E-10	6.59E-11
SCO22085SKAG	4.27E-12	1.96E-10	2.45E-17	2.94E-13	1	1.20E-14	3.12E-11	1.85E-11	2.36E-12	4.80E-16	3.48E-10	5.76E-13
SCO22086SKAG	4.55E-07	0.0001903	8.93E-10	3.92E-08	0.9964766	1.87E-08	1.33E-06	0.0033285	9.98E-07	9.53E-09	2.76E-08	1.73E-06

	1	2	3	4	5	6	7	8	9	10	11	12
SCO22087SKAG	1.68E-06	1.51E-05	2.95E-09	4.59E-09	0.8951239	8.13E-10	6.76E-05	2.31E-06	7.93E-06	0.00080282	0.1039733	5.33E-06
SCO22096SKAG	3.81E-05	2.26E-07	1.34E-07	8.42E-09	1.04E-08	3.49E-10	0.000184	8.16E-08	1.82E-05	1.58E-10	0.9997592	5.89E-08
SCO22097SKAG	0.0006813	0.8925378	0.0003598	0.009608	0.0047785	0.0035008	0.0344295	0.0227651	0.01807862	0.00015713	1.73E-05	0.01308629
SCO22098SKAG	0.0003546	0.9330015	0.0002831	0.0006389	0.0010392	0.0001909	0.0010835	0.0402266	0.02141908	2.34E-05	0.00114129	0.00059801
SCO22099SKAG	9.68E-09	1.97E-06	1.37E-08	1.25E-08	0.2265901	3.96E-09	7.50E-05	0.7700313	0.00167588	2.46E-09	0.0016257	1.05E-08
SCO22100SKAG	0.9135705	0.0013939	2.23E-09	2.72E-06	0.0847421	3.75E-08	1.28E-06	0.0001796	3.31E-05	4.33E-08	7.56E-05	9.27E-07
SCO22101SKAG	2.35E-07	0.0040411	5.52E-06	0.0381724	1.06E-05	6.54E-05	0.0124737	0.9252301	0.01749025	1.44E-06	0.00250778	1.44E-06
SCO22102SKAG	1	7.72E-09	1.69E-14	2.12E-09	8.63E-12	2.12E-12	6.56E-12	1.55E-10	4.58E-12	1.31E-12	9.90E-12	3.95E-11
SCO22104SKAG	0.0006481	0.0157505	1.70E-07	6.87E-06	0.0006921	9.01E-07	1.16E-06	0.9823872	0.00037467	3.52E-07	1.55E-06	0.00013649
SCO22105SKAG	0.9940669	9.15E-06	7.15E-11	1.39E-08	2.55E-07	7.46E-10	1.12E-08	1.32E-06	8.94E-09	0.00591075	1.13E-05	2.24E-07
SCO22106SKAG	1.67E-11	2.26E-08	4.79E-15	1.27E-08	1	8.99E-12	2.79E-09	3.32E-10	1.66E-09	2.21E-13	2.78E-11	8.98E-10
SCO22108SKAG	3.80E-07	7.53E-06	1.36E-11	0.072977	1.29E-08	6.76E-09	1.27E-09	0.9268806	0.00013371	1.31E-09	7.43E-07	7.58E-11
SCO22110SKAG	0.0003253	0.0001383	7.68E-09	3.71E-07	1.28E-05	1.42E-07	3.01E-07	0.6493129	1.27E-06	0.34956895	0.00063028	9.36E-06
SCO22111SKAG	0.9987041	0.0012874	1.07E-08	1.40E-06	3.89E-08	3.10E-08	6.83E-09	6.77E-07	4.97E-06	1.92E-07	5.21E-07	5.82E-07
SCO22112SKAG	2.57E-07	9.57E-05	1.66E-08	6.03E-08	0.9957208	0.0041066	2.07E-06	4.05E-05	1.07E-05	4.58E-09	1.03E-05	1.29E-05
SCO22113SKAG	0.0289304	3.03E-05	1.54E-08	4.58E-07	6.71E-07	2.14E-08	2.49E-08	0.971004	1.95E-05	2.76E-07	1.43E-05	9.34E-10
SCO22114SKAG	6.68E-12	2.09E-06	0.999993	1.80E-10	5.70E-12	8.89E-10	1.91E-12	3.18E-08	4.84E-06	4.67E-09	2.00E-09	1.41E-10
SCO22115SKAG	4.27E-08	9.95E-05	1.06E-08	0.0054556	0.9549384	2.02E-08	9.93E-05	0.0218047	0.01754088	2.15E-08	6.15E-05	1.13E-07
SCO22116SKAG	4.50E-07	0.0013891	2.38E-06	1.52E-07	2.02E-07	0.9234183	1.59E-06	0.0750976	2.51E-06	4.36E-06	1.73E-05	6.61E-05
SCO22117SKAG	7.59E-07	0.004091	0.0014986	1.43E-05	0.9939798	2.32E-05	3.99E-05	9.24E-05	7.87E-05	5.96E-07	8.39E-05	9.68E-05
SCO22178SKAG	5.18E-07	4.53E-05	6.40E-06	3.09E-07	1.25E-05	0.1201093	6.85E-05	0.0039998	5.28E-05	7.06E-08	4.76E-06	0.87569974
SCO22186SKAG	1.28E-06	0.000821	6.77E-05	5.76E-07	2.96E-05	0.2596113	0.01441	0.0019366	0.66868636	4.26E-07	0.04917751	0.0052576
SCO22188SKAG	0.0056951	0.780472	2.60E-08	0.0048466	0.0002754	4.89E-06	0.0005807	0.001544	0.20637319	3.34E-06	2.00E-05	0.00018473
SCO22192SKAG	0.9498917	0.0482027	1.73E-06	3.42E-06	2.67E-06	6.42E-07	0.0001739	1.57E-05	0.00118408	8.30E-07	0.00034562	0.000177
SCO22193SKAG	1	1.83E-08	1.09E-14	2.76E-09	5.17E-13	1.02E-11	1.16E-10	4.91E-11	1.93E-10	5.57E-11	9.81E-11	2.24E-10
SCO22129SKAG	4.56E-07	0.0002077	8.99E-09	1.21E-09	5.90E-09	2.94E-08	5.89E-07	7.95E-08	1.85E-06	0.999787	3.17E-07	2.01E-06
SCO22134SKAG	3.61E-08	0.002008	1.04E-05	5.44E-06	2.65E-07	1.29E-07	0.0428326	2.53E-06	0.95481387	1.69E-08	2.82E-05	0.00029866
SCO22136SKAG	1.05E-05	0.9411833	1.80E-05	0.0005418	4.13E-06	2.15E-06	0.0012619	5.85E-07	0.05093856	6.68E-06	2.68E-08	0.00603238
SCO22139SKAG	9.11E-10	3.41E-07	3.64E-12	8.64E-10	3.86E-10	4.06E-10	4.87E-10	0.9999996	9.38E-09	1.24E-10	3.34E-09	1.10E-10
SCO22205SKAG	0.0002486	0.9916816	0.0009071	0.0005355	7.43E-06	0.0002051	3.40E-05	4.80E-05	0.00595495	7.99E-05	3.44E-06	0.00029448
SCO22168SKAG	2.54E-08	0.00953	0.0001055	4.37E-07	0.947384	7.34E-07	0.0412586	8.70E-08	0.00139207	1.04E-07	3.20E-05	0.0002964
19PR0184	3.76E-05	0.4237264	0.0010151	0.001244	4.79E-06	0.0008106	0.0003863	7.34E-06	0.00096847	6.91E-05	3.98E-08	0.57173021
19PR0185	3.78E-06	0.0024043	1.24E-07	0.0030706	4.02E-07	4.35E-07	3.71E-08	0.9861176	0.00839984	6.24E-07	2.28E-06	1.35E-08
19PR0187	2.13E-13	8.31E-11	1.09E-12	1.10E-13	6.17E-13	1.06E-13	1.72E-09	2.46E-12	1.48E-11	0.99999999	1.01E-08	8.78E-12
19PR0188	6.76E-07	0.0137207	0.0061618	1.47E-07	5.13E-05	9.84E-06	0.0045808	2.92E-07	0.00125819	8.71E-07	9.66E-07	0.97421441
19PR0191	5.03E-06	0.0761428	0.1726858	0.0001082	3.41E-05	5.48E-05	0.0089111	0.0004419	0.73424658	5.16E-05	0.00297089	0.0043472
19PR0194	0.0016936	0.0059118	1.03E-05	5.60E-05	0.5186998	4.90E-06	0.4627638	2.23E-05	0.00669723	8.43E-07	0.00398425	0.00015522
19PR0196	2.00E-06	0.009182	4.27E-05	0.1084051	1.46E-05	3.23E-06	0.0006842	0.0001231	0.88077626	1.29E-07	0.0007652	1.51E-06
19PR0197	5.86E-07	0.0061482	2.27E-07	0.99316	4.05E-07	1.72E-05	1.75E-05	2.05E-06	0.00063281	6.14E-08	5.56E-09	2.10E-05
19PR0198	2.81E-07	0.0002557	0.0005693	1.13E-08	3.99E-06	0.0080936	2.68E-05	1.53E-07	0.00011988	4.56E-09	2.12E-06	0.99092812
19PR0200	4.51E-05	5.31E-07	0.0001975	2.26E-09	4.59E-07	0.0001174	0.0002697	1.36E-05	0.00821127	2.98E-09	0.99112215	2.22E-05
19PR0253	1.38E-05	0.0027015	0.0001573	0.0014352	7.81E-05	3.03E-06	0.0007253	0.0142938	0.64105547	4.54E-07	0.33953446	1.58E-06
19PR0254	6.07E-06	0.0250576	5.99E-06	5.88E-05	8.22E-07	2.90E-06	0.0001311	8.36E-05	0.00019703	0.97436237	3.91E-06	8.97E-05
19PR0255	1.99E-06	0.0104076	0.0003947	0.0581656	4.72E-06	0.4485226	0.0002999	0.0001101	0.46945744	6.81E-07	0.00022677	0.01240799
19QK0852	0.0065174	0.0030759	1.51E-08	0.990367	1.52E-09	2.02E-05	5.14E-06	6.53E-07	1.02E-05	3.06E-06	1.75E-08	3.93E-07
19QK0854	2.55E-07	0.0109354	0.028022	6.07E-07	3.68E-05	3.77E-07	0.2302792	1.82E-06	0.71527827	2.21E-07	0.01316909	0.00227589
19QK0877	6.03E-14	1.13E-12	1.17E-16	8.01E-13	1.87E-16	7.08E-15	4.60E-15	1.06E-14	6.38E-15	1	2.58E-12	1.58E-14
19QK0885	2.40E-09	5.48E-06	6.51E-07	4.20E-06	2.54E-09	7.36E-08	1.77E-05	0.0023201	2.83E-05	0.99630874	0.0013147	2.03E-08
19QK0911	5.98E-05	0.8493895	0.0001763	0.0900677	1.93E-05	0.0112192	0.0081463	0.0168399	0.02381693	5.01E-05	1.30E-05	0.00020194
19QK0915	0.0427033	0.0251414	6.30E-06	7.66E-05	2.03E-08	0.931751	5.58E-06	4.40E-05	4.54E-05	3.92E-06	1.28E-05	0.00020975

	1	2	3	4	5	6	7	8	9	10	11	12
19QK1381	7.70E-07	0.0872681	3.32E-08	0.8091754	3.22E-08	0.0003	0.0004423	0.1022097	0.000595	5.69E-06	9.60E-07	2.00E-06
SCO22232SKAG	0.000195	0.9856236	2.07E-06	0.0005298	1.54E-05	5.56E-06	2.01E-05	1.48E-05	0.01306094	2.37E-06	1.85E-07	0.00053015
SCO22261SKAG	6.10E-05	0.014996	3.37E-06	0.0007126	9.14E-06	6.56E-06	0.0001107	8.14E-06	0.00016586	0.9813199	2.75E-05	0.00257928
SCO22263SKAG	1.70E-08	0.0157734	9.90E-05	0.9106673	2.93E-07	2.94E-05	7.44E-05	3.70E-05	0.07330494	5.73E-07	4.03E-08	1.37E-05
SCO22270SKAG	5.05E-09	0.0001011	4.21E-06	4.70E-05	1.27E-06	5.20E-07	0.0104873	0.0001807	0.9890369	3.78E-08	0.00013356	7.33E-06
SCO22271SKAG	4.38E-09	2.94E-05	0.0001059	2.92E-08	8.66E-09	0.0404776	6.04E-05	1.66E-08	6.26E-06	1.89E-08	8.26E-08	0.95932029
SCO22276SKAG	6.10E-08	0.0010593	0.9775053	1.01E-07	3.83E-07	4.04E-07	6.87E-05	5.18E-07	0.00369826	1.14E-06	4.89E-06	0.01766095
SCO22277SKAG	3.89E-08	0.0005873	1.27E-05	0.9875087	5.64E-07	5.40E-05	0.0041034	5.01E-07	0.00687957	2.12E-08	2.41E-07	0.00085299
SCO22470SKAG	8.13E-10	4.61E-06	3.30E-08	0.0141295	0.1103023	4.20E-09	0.0123815	0.026097	0.83026442	1.24E-09	0.00682036	1.82E-07
SCO22717SKAG	5.72E-06	0.0076738	1.36E-08	0.512558	0.3042377	3.76E-06	3.81E-06	0.1706544	0.00484384	1.48E-05	3.39E-06	8.28E-07
SCO22876SKAG	0.0018333	1.06E-05	5.15E-07	3.72E-09	2.59E-06	3.45E-09	0.9541492	3.97E-08	0.00016582	0.02607786	0.01474521	0.00301476
SCO22889SKAG	6.38E-09	0.0016373	0.6800725	2.89E-06	1.01E-06	5.77E-05	0.001869	3.01E-08	0.00027532	9.65E-07	6.53E-08	0.31608324
SCO22940SKAG	1.20E-08	6.31E-05	7.07E-06	4.24E-06	1.04E-09	7.47E-08	7.29E-07	1.06E-08	3.94E-06	0.99991364	8.67E-08	7.13E-06
SCO22941SKAG	9.48E-08	0.0004541	1.33E-06	9.31E-08	1.57E-07	0.0032368	0.0037753	1.42E-07	0.00261556	3.01E-10	7.91E-05	0.98983721
SCO22945SKAG	1.71E-06	0.9981628	2.32E-06	0.0001751	3.32E-07	1.48E-05	0.0003724	6.60E-06	0.00120512	1.01E-07	6.05E-07	5.81E-05
SCO22947SKAG	9.84E-09	0.0001039	6.69E-06	9.25E-06	2.36E-07	4.82E-06	0.0021192	2.21E-07	3.98E-05	0.99731579	6.51E-06	0.00039366
SCO22547SKAG	3.76E-09	0.0001929	3.86E-10	2.12E-06	3.58E-09	2.17E-06	0.9905243	1.10E-08	1.50E-06	0.0090239	0.00023098	2.20E-05
SCO22551SKAG	9.94E-09	9.61E-05	7.19E-06	4.78E-08	1.08E-09	3.65E-07	0.0001704	1.08E-08	1.54E-06	0.99959979	0.00010414	2.04E-05
SCO22553SKAG	2.06E-09	0.0005368	0.9988049	8.28E-08	2.88E-09	2.03E-07	2.54E-06	1.43E-08	0.00055951	2.01E-07	4.15E-07	9.54E-05
SCO22580SKAG	1.52E-05	0.0172162	0.0069325	0.3452936	2.62E-05	0.0006806	0.0020587	8.49E-05	0.60712941	8.79E-05	0.00044467	0.02003012
SCO22588SKAG	0.0001203	0.0001393	0.0004355	3.14E-08	1.13E-05	1.12E-07	0.9511412	3.78E-07	0.00100562	2.21E-08	0.04686625	0.00028004
SCO22599SKAG	2.85E-09	0.000482	8.75E-08	4.34E-07	4.56E-08	0.9754675	0.0002718	1.03E-07	1.13E-06	2.61E-08	4.90E-10	0.02377684
SCO22607SKAG	0.9990709	0.0008891	6.36E-08	1.82E-06	1.86E-07	2.84E-08	7.72E-08	1.35E-06	7.12E-06	4.33E-08	1.45E-05	1.49E-05
19PR0226	1.43E-05	0.012333	4.42E-05	1.83E-05	1.19E-05	0.8077164	0.0003294	0.0001334	0.00209702	1.39E-07	4.10E-05	0.17726088
19PR0229	1.46E-05	0.1547494	2.84E-05	0.0129969	8.10E-06	1.57E-05	0.0195516	0.0001535	0.81144704	3.90E-07	0.00017293	0.00086149
19PR0230	5.67E-07	0.0034957	1.74E-06	2.30E-05	0.3111706	0.5536582	0.070353	2.68E-05	0.00312955	4.18E-08	0.02975736	0.02838343
19PR0231	1.12E-10	1.96E-05	0.990794	1.89E-07	5.56E-09	6.21E-07	0.0011677	1.04E-08	0.00736735	3.76E-08	7.72E-05	0.00057337
19PR0233	1.06E-08	2.09E-05	6.49E-06	7.49E-09	3.23E-08	0.0129801	0.000154	3.12E-09	5.38E-07	0.01999782	1.20E-06	0.96683892
19PR0235	1.29E-09	8.30E-05	0.000418	4.70E-07	4.56E-10	1.12E-07	1.49E-08	1.50E-09	5.68E-07	0.99948632	2.40E-09	1.15E-05
19PR0236	4.07E-09	0.0012978	0.3909623	8.81E-09	1.21E-08	0.0804144	6.41E-05	3.94E-08	0.00078967	1.12E-07	3.87E-06	0.52646775
19PR0241	1.71E-09	7.53E-06	1.78E-05	2.69E-09	4.50E-11	3.68E-08	1.54E-08	7.20E-08	4.32E-07	0.99997403	9.51E-08	1.60E-08
19PR0244	3.23E-06	0.2006214	0.7948448	2.94E-06	7.82E-06	5.58E-05	4.15E-06	6.52E-05	0.003277	2.23E-05	7.37E-07	0.00109456
19PR0245	4.64E-15	1.10E-08	3.82E-11	2.13E-12	2.92E-13	0.9999852	1.09E-06	9.25E-12	8.74E-10	5.75E-14	4.03E-10	1.37E-05
19PR0246	1.42E-07	0.1728759	0.8258716	8.12E-06	1.21E-08	0.000362	6.88E-06	8.48E-08	3.80E-05	0.0006018	1.26E-06	0.00023424
19PR0247	8.50E-08	2.78E-05	8.91E-05	6.94E-08	1.34E-06	0.0464198	1.86E-05	9.78E-08	1.80E-05	2.58E-08	8.74E-07	0.95342432
19QK1067	4.70E-12	5.33E-06	2.80E-07	2.09E-09	2.01E-09	8.11E-08	0.9982649	5.11E-09	0.00168247	2.61E-10	4.10E-05	5.90E-06
19QK1070	1.80E-12	6.90E-08	0.9999861	1.57E-10	6.93E-12	1.23E-10	4.59E-11	6.58E-10	1.38E-05	2.50E-10	2.03E-08	1.33E-08
19QK1073	2.55E-11	3.38E-08	7.35E-09	6.96E-08	7.67E-09	1.86E-10	0.0005781	0.0011315	0.00012862	0.00033331	0.99782829	2.13E-10
19QK1074	5.32E-10	0.0001307	0.0004051	5.31E-08	3.37E-07	1.95E-09	0.0168668	2.12E-08	0.98236862	2.73E-10	0.00021066	1.77E-05
19QK1076	2.52E-08	0.0352538	0.0276222	0.0020575	0.008092	5.97E-06	0.0049111	2.68E-06	0.92139707	1.80E-06	0.0004408	0.00021496
19QK1078	2.19E-09	5.75E-05	8.04E-09	0.9945345	1.48E-07	1.05E-06	0.0001799	0.0025424	0.00268287	1.75E-08	1.56E-06	8.77E-08
19QK1084	5.36E-08	0.0121311	0.0214474	0.0001274	1.87E-07	0.1509257	0.0002066	7.94E-06	0.80702409	1.59E-06	0.00054128	0.00758666
19QK1091	8.38E-07	0.0075505	2.10E-05	0.0002393	1.13E-05	8.45E-05	0.0004668	0.9874355	0.00416944	3.15E-06	1.01E-05	7.50E-06
19QK1095	1.50E-05	0.8392087	0.0850907	6.26E-05	1.60E-05	0.0028946	0.001244	7.41E-05	0.00198719	0.00110976	9.85E-06	0.06828739
19QK1097	3.73E-06	0.3761963	1.20E-05	1.42E-05	0.5755873	0.0002116	0.0270288	6.32E-05	0.0005733	6.18E-06	1.08E-06	0.02030242
19QK1104	0.0004434	0.9757729	0.000271	0.0020977	0.0017524	0.0012087	0.0014025	0.0043222	0.01074605	2.17E-05	1.65E-05	0.00194487
19QK1106	2.53E-11	8.59E-06	0.8713044	9.67E-10	1.41E-10	0.0719674	3.05E-06	1.22E-08	1.05E-05	1.49E-07	8.11E-08	0.0567059
19QK1108	8.49E-11	0.0001774	4.18E-08	4.00E-06	3.00E-10	0.9997291	1.88E-05	2.50E-08	1.39E-07	6.89E-09	1.42E-10	7.04E-05
19QK1112	4.55E-07	0.0032933	1.18E-06	0.0917365	4.15E-09	0.9046752	1.29E-07	1.60E-06	0.00019933	2.97E-06	4.61E-06	8.47E-05
19QK1135	4.75E-09	0.00787	1.88E-05	1.60E-06	2.82E-09	0.9845531	0.0001778	2.81E-07	0.00022555	1.04E-06	2.65E-08	0.00715166

	1	2	3	4	5	6	7	8	9	10	11	12
SCO221005SKAG	9.38E-07	0.0939026	0.0017949	0.0083499	1.29E-05	0.0007719	0.7491311	3.51E-05	0.11022874	2.99E-06	1.64E-05	0.03575263
SCO221006SKAG	0.0002086	0.9580732	0.0004459	0.0221498	2.29E-05	0.0031699	0.0022669	0.0008413	0.00592495	2.03E-05	0.00011205	0.00676416
SCO221008SKAG	2.47E-09	0.0260536	0.0044613	6.40E-07	2.87E-08	0.9146124	0.0176652	1.21E-06	0.00548561	9.80E-08	8.08E-05	0.03163905
SCO221010SKAG	1.65E-07	0.9893283	8.22E-08	3.17E-05	3.05E-08	0.000276	0.0103555	7.06E-07	4.08E-06	3.29E-07	1.59E-07	3.01E-06
SCO221011SKAG	7.37E-08	0.2690479	0.0005073	0.6029261	2.24E-06	0.0005437	0.0464877	7.05E-06	0.0795158	1.44E-05	1.11E-06	0.00094665
SCO221023SKAG	1.31E-11	5.50E-06	0.9902952	1.09E-07	6.95E-11	0.0031432	1.58E-05	7.44E-09	0.00552317	1.37E-08	7.80E-05	0.00093907
SCO221028SKAG	3.13E-08	0.3234517	0.3563392	1.80E-06	5.32E-06	2.04E-05	0.1539964	2.81E-06	0.16502306	7.74E-06	3.81E-05	0.00111351
SCO221035SKAG	7.18E-05	0.0001897	0.0003327	1.72E-07	1.42E-07	6.09E-08	0.989448	6.25E-07	0.00832433	1.69E-08	0.00074183	0.0008906
SCO221040SKAG	4.27E-10	5.21E-05	2.29E-06	5.32E-08	5.11E-08	0.0012146	0.7469346	2.65E-08	0.01853149	9.14E-11	1.49E-05	0.23324989
SCO221041SKAG	3.78E-06	0.0053662	6.45E-05	4.85E-07	9.83E-05	0.0554222	0.0378185	0.0001586	0.63378081	1.87E-08	0.03806513	0.2292215
SCO221053SKAG	0.0001935	0.7242694	0.0016988	0.0030533	1.91E-05	2.95E-05	8.23E-06	0.0006336	0.2698059	6.05E-05	5.20E-05	0.00017621
SCO221064SKAG	2.21E-06	0.9662967	0.0234846	0.0001725	1.10E-07	0.0056901	0.0002273	2.68E-05	0.00121551	0.00229982	4.22E-07	0.00058389
SCO221066SKAG	2.93E-10	0.0001552	0.000129	7.47E-09	5.45E-06	6.14E-08	0.9970047	6.73E-09	0.00180514	9.32E-10	5.58E-06	0.00089487
SCO221068SKAG	1.19E-07	0.0017582	0.000168	4.08E-07	9.93E-07	0.9165929	0.0005314	5.69E-06	0.00018875	1.02E-06	2.31E-06	0.08075031
SCO221069SKAG	1.41E-15	1.94E-08	0.9999994	1.97E-14	4.30E-14	1.91E-12	2.92E-11	1.70E-14	4.53E-09	3.39E-12	1.41E-13	5.40E-07
SCO221070SKAG	5.10E-08	0.0049962	6.53E-06	0.0002446	2.31E-07	0.3910525	0.0029925	1.37E-07	0.00523368	3.34E-08	1.10E-06	0.59547251
SCO221071SKAG	6.96E-07	0.5280803	0.0536995	0.0848771	3.78E-07	0.0007918	0.0059622	6.75E-05	0.32482054	0.00117441	0.00013451	0.00039115
SCO221074SKAG	0.0034419	0.4406433	2.38E-06	8.74E-06	0.000322	3.18E-05	0.0010838	0.2731719	0.00652832	0.00012227	2.66E-06	0.27464094
SCO221076SKAG	2.24E-09	0.006658	3.66E-05	4.48E-05	1.21E-07	3.31E-05	0.9904438	4.23E-06	0.00273105	8.35E-08	2.64E-05	2.18E-05
SCO221077SKAG	0.9589673	0.0407662	6.71E-06	4.56E-06	7.05E-05	1.28E-07	7.72E-06	3.95E-06	0.00012467	2.53E-07	8.15E-07	4.72E-05
SCO221079SKAG	8.84E-09	0.0025046	0.0011704	9.53E-07	6.33E-08	0.9840686	0.0028963	3.38E-06	0.00213489	2.93E-07	4.41E-05	0.0071764
SCO221081SKAG	1.80E-07	0.0036613	6.78E-08	2.81E-05	9.96E-10	0.9905797	6.21E-06	3.74E-08	1.76E-05	5.93E-07	1.06E-07	0.00570604
SCO221083SKAG	1.73E-06	0.8928407	0.0067719	0.0001141	2.41E-06	1.73E-05	0.0091551	4.78E-06	0.0175344	2.38E-07	1.14E-05	0.07354603
SCO221084SKAG	1.62E-09	0.0001041	7.29E-07	6.76E-06	1.15E-08	0.9709663	9.98E-06	0.0288286	9.40E-06	2.41E-07	2.36E-08	7.38E-05
SCO221086SKAG	1.18E-06	0.9515801	6.16E-06	0.0001926	1.01E-07	7.48E-05	0.0285879	9.66E-06	0.01880274	7.26E-07	0.00073374	1.03E-05
SCO221087SKAG	2.73E-09	0.0001041	7.57E-05	1.82E-05	2.50E-09	0.4961233	2.10E-05	1.52E-08	0.00025176	2.28E-07	1.12E-07	0.50340546
SCO221089SKAG	4.45E-08	6.69E-05	3.55E-05	2.37E-08	6.47E-07	0.0106766	7.29E-05	4.92E-07	5.71E-05	3.11E-09	1.31E-06	0.98908835
SCO221090SKAG	1.90E-07	0.0965483	0.0644563	0.0007513	8.49E-06	0.0024789	0.7956487	6.31E-05	0.0311595	1.08E-05	0.00750629	0.00136804
SCO221091SKAG	1.39E-06	0.9994785	1.61E-06	7.76E-08	1.61E-07	2.67E-06	0.0002666	5.57E-07	0.00012768	6.72E-07	8.90E-08	0.00011999
SCO221093SKAG	3.46E-09	0.0014613	3.14E-07	0.0001118	1.84E-08	0.9611232	0.0244818	1.66E-07	0.00185533	4.43E-08	9.18E-07	0.01096512
SCO221094SKAG	1.37E-05	0.0042662	2.47E-10	0.000152	1.55E-07	1.49E-06	0.0002457	1.05E-05	0.00023868	0.99501141	1.04E-06	5.90E-05
SCO221095SKAG	9.17E-11	4.80E-05	0.9982575	3.03E-08	1.43E-08	7.65E-07	3.63E-05	1.15E-06	0.00164608	1.22E-06	8.30E-06	5.83E-07
SCO221096SKAG	4.91E-13	6.14E-06	2.62E-09	1.43E-09	3.61E-11	1.95E-09	0.9993632	1.81E-10	0.00062982	2.77E-12	2.55E-07	5.71E-07
SCO221098SKAG	0.0008666	0.0001568	8.32E-08	2.60E-06	5.23E-09	3.30E-07	1.44E-06	1.48E-07	5.68E-08	0.9989716	2.58E-07	4.69E-08
SCO221107SKAG	5.71E-08	3.60E-05	1.99E-07	1.49E-07	0.8407217	8.46E-08	0.0015134	0.1569524	0.00059451	2.13E-08	0.00017036	1.12E-05
SCO221109SKAG	5.80E-09	0.0012888	8.16E-11	0.9985588	9.91E-10	3.01E-06	1.68E-06	4.39E-08	0.00014746	1.26E-08	1.09E-10	1.05E-07
SCO221110SKAG	2.43E-08	0.0019925	3.64E-07	8.29E-07	2.20E-08	0.6702288	0.0003047	1.83E-06	0.00014281	7.35E-08	7.22E-09	0.32732812
SCO221111SKAG	1.21E-12	2.71E-06	0.9999424	1.83E-09	3.51E-11	4.99E-09	2.91E-06	2.14E-10	4.74E-05	3.85E-09	1.85E-06	2.67E-06
SCO221113SKAG	1.87E-10	0.0001019	0.9978562	1.91E-09	3.86E-09	4.88E-08	2.24E-06	9.95E-07	0.00203297	3.35E-08	5.28E-06	2.78E-07
SCO221117SKAG	6.77E-12	1.36E-05	3.77E-08	4.40E-08	8.19E-05	0.0002903	0.9711063	1.22E-08	0.00930633	4.00E-11	0.01871399	0.00048751
SCO221118SKAG	4.24E-06	0.0152869	1.91E-05	1.60E-06	1.13E-05	2.00E-06	0.0091736	0.239694	0.73330345	7.20E-06	0.00055084	0.0019457
SCO221119SKAG	5.39E-08	1.51E-05	7.35E-08	3.23E-06	1.10E-08	0.0412837	9.72E-08	3.56E-06	4.57E-07	0.95861887	7.13E-05	3.54E-06
SCO221120SKAG	3.67E-07	0.0011249	0.4268232	1.30E-07	2.30E-06	0.5676847	3.43E-07	0.0017485	0.00034146	2.52E-06	0.00216751	0.00010409
SCO221121SKAG	3.35E-08	0.000117	1.95E-09	0.000989	3.20E-08	6.20E-07	1.12E-07	0.9988883	3.78E-06	1.10E-07	1.44E-10	1.06E-06
SCO221122SKAG	7.67E-10	2.02E-06	3.83E-05	1.61E-05	6.24E-08	2.41E-10	4.32E-05	2.67E-07	0.99962038	1.95E-11	0.0002783	1.36E-06
SCO221123SKAG	1.59E-11	9.28E-06	4.97E-05	3.17E-09	3.09E-08	7.62E-08	0.9930891	3.17E-08	0.00493102	7.18E-10	0.00161803	0.00030265
SCO221124SKAG	4.23E-07	0.0061176	1.61E-07	0.0011789	1.08E-06	0.0001649	0.0103587	0.9812054	0.00093431	1.39E-06	3.43E-05	2.85E-06
SCO221125SKAG	2.69E-08	8.14E-05	3.21E-07	2.61E-09	1.38E-06	0.0029985	0.0023198	6.90E-08	0.00014531	1.66E-10	6.54E-07	0.99445261
SCO221128SKAG	1.42E-10	0.0001191	1.11E-05	1.99E-07	8.91E-09	1.66E-07	0.9800782	8.54E-09	0.00042113	0.01872187	0.00056239	8.58E-05
SCO221129SKAG	0.0001277	0.8303125	2.28E-05	0.0379868	4.15E-06	0.0028396	0.0004989	0.0869293	0.00728501	0.00603413	2.82E-06	0.02795625

	1	2	3	4	5	6	7	8	9	10	11	12
SCO221132SKAG	1.07E-06	0.0039721	0.0002137	2.31E-07	1.39E-06	2.30E-06	0.0194716	6.08E-07	0.00032977	0.8391034	2.54E-05	0.1368785
SCO221135SKAG	2.29E-08	0.0215752	0.0001187	0.0002771	7.50E-07	1.09E-06	0.4182282	7.04E-07	0.55370809	1.02E-07	0.00010373	0.00598634
SCO22964SKAG	7.78E-07	0.9867382	5.05E-07	8.66E-05	2.16E-06	0.0005721	0.0109843	2.04E-06	3.22E-05	9.63E-07	3.14E-08	0.00158019
SCO22965SKAG	3.95E-07	0.0022453	7.33E-07	6.12E-06	5.52E-06	1.14E-05	0.2823954	4.08E-07	0.00023033	0.71089755	0.000366	0.0038409
SCO22966SKAG	6.32E-06	0.9844997	0.0009559	8.39E-05	1.15E-05	0.0004937	0.0080647	0.0001072	0.00049691	2.05E-05	5.11E-06	0.00525464
SCO22967SKAG	3.51E-10	0.0003537	1.95E-07	3.17E-07	2.10E-08	3.42E-08	0.8590687	0.0010746	0.13946349	4.54E-10	4.07E-06	3.48E-05
SCO22969SKAG	1.43E-06	0.8886401	1.57E-06	2.05E-05	1.94E-05	0.0002105	0.1096169	0.0001086	0.00050301	6.35E-07	2.63E-06	0.00087466
SCO22970SKAG	1.81E-07	0.0001712	1.86E-06	2.11E-05	2.76E-08	1.52E-07	8.17E-06	2.05E-07	0.00064575	0.99905538	7.73E-05	1.87E-05
SCO22974SKAG	3.16E-08	0.0013106	1.00E-07	4.68E-07	1.12E-07	0.975892	2.48E-05	0.0048959	5.31E-07	5.44E-08	1.25E-09	0.01787538
SCO22977SKAG	2.11E-06	0.0075987	1.01E-05	0.0003215	2.02E-06	0.981763	3.08E-05	0.0006963	0.00030114	2.99E-07	3.36E-05	0.00924049
SCO22978SKAG	7.07E-10	0.0040521	0.995847	6.76E-08	8.43E-10	1.77E-06	1.00E-07	3.22E-07	9.68E-05	1.15E-06	6.01E-09	6.97E-07
SCO22980SKAG	1.26E-10	0.0012709	0.6352158	3.56E-09	5.36E-10	0.1056103	1.36E-05	1.65E-09	2.79E-05	5.08E-08	5.04E-08	0.25786138
SCO22981SKAG	6.16E-09	0.0049855	0.0039461	0.0009316	5.04E-09	0.963312	3.60E-07	9.52E-08	2.80E-05	1.97E-07	8.65E-10	0.02679608
SCO22987SKAG	2.43E-05	0.9819779	0.0068944	9.40E-05	2.16E-06	0.0005228	0.0007258	0.0002902	0.00768679	0.00049381	1.34E-06	0.00128652
SCO22999SKAG	6.65E-06	0.0294632	0.0001235	0.0857508	3.80E-05	2.57E-06	4.01E-06	0.0022577	0.88233122	8.21E-07	2.10E-05	5.64E-07

RESERVOIR NATIVE FISH BASELINE GENETICS STUDY
TECHNICAL MEMORANDUM FOR EXPERT PANEL

ATTACHMENT D

POSTERIOR ASSIGNMENTS OF BULL TROUT AND DOLLY VARDEN

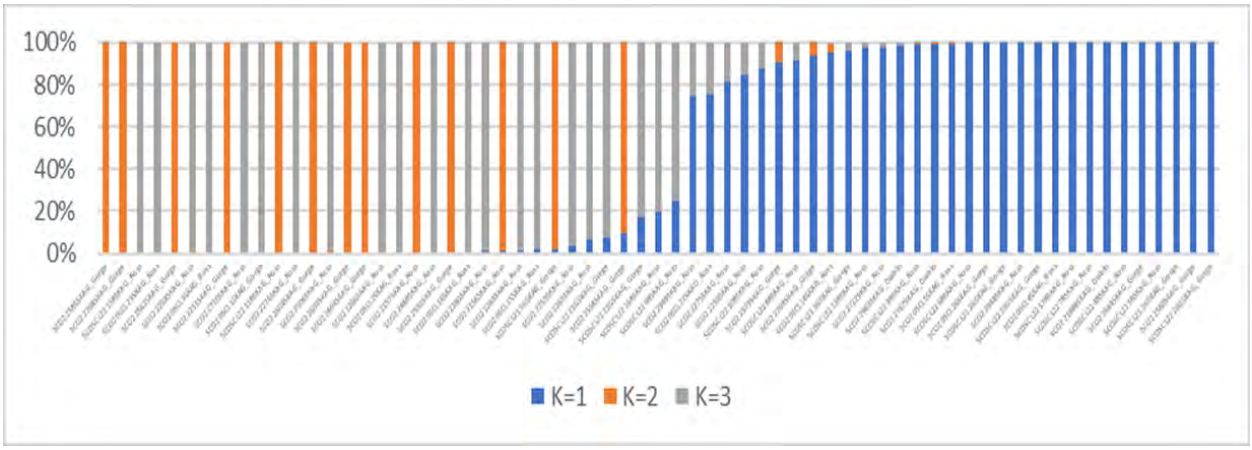


Figure D-1. Posterior assignments of Bull Trout Sorted by *k*=3.

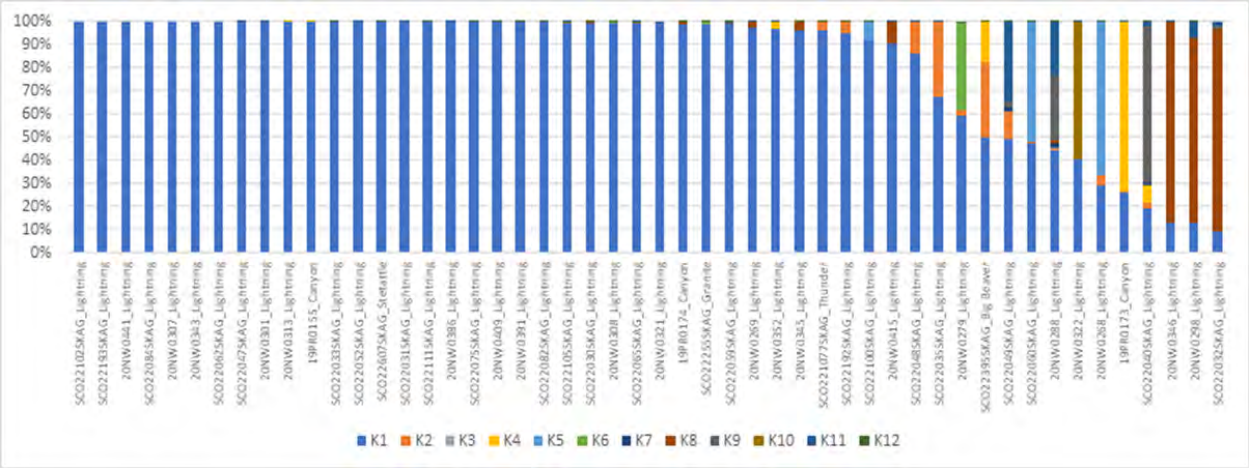


Figure D-2. Posterior assignments of Dolly Varden sorted by *k*=1.

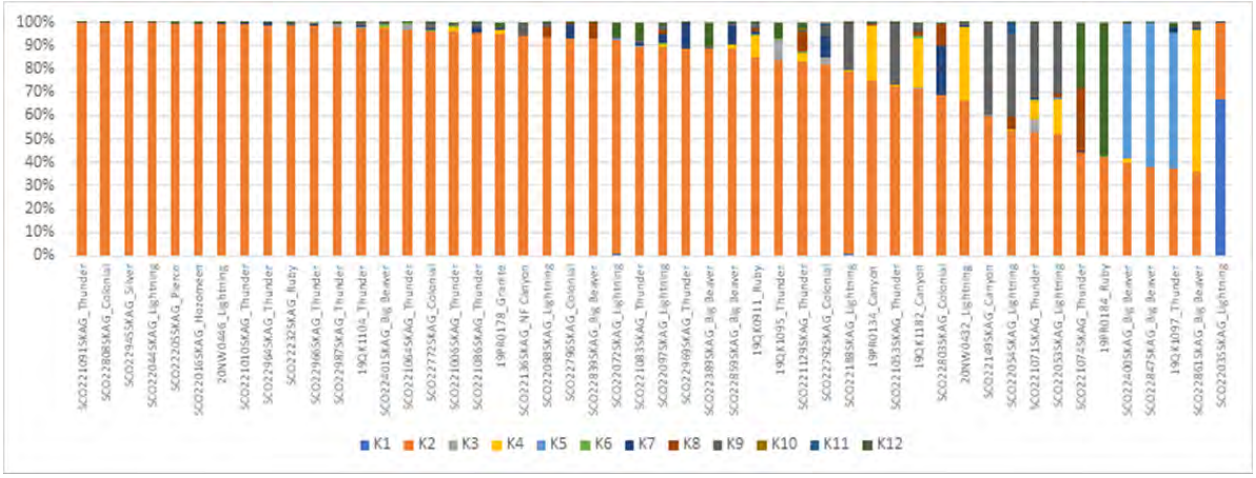


Figure D-3. Posterior sorted by *k*=2.

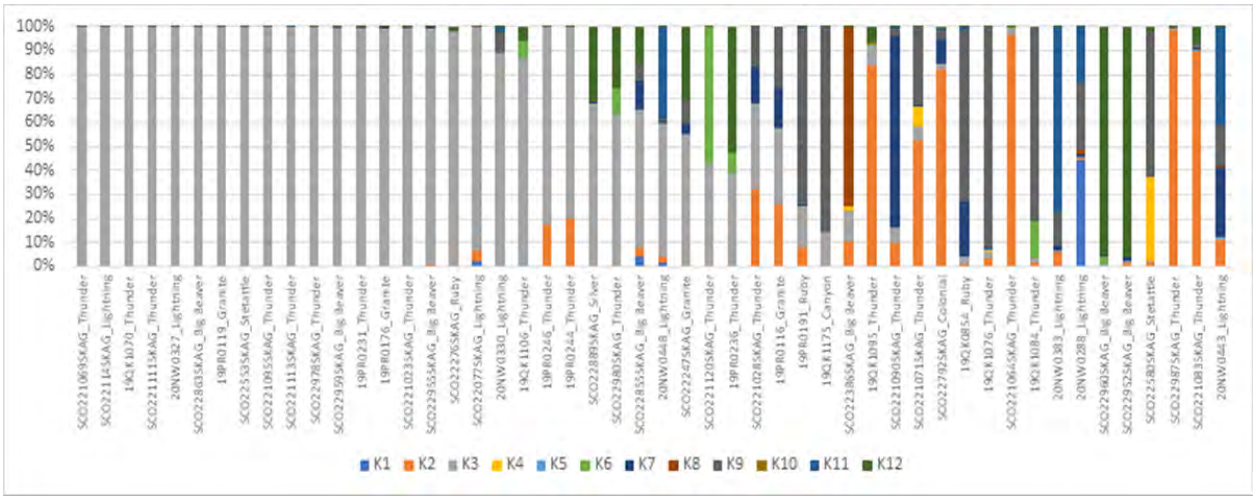


Figure D-4. Posterior sorted by *k*=3.

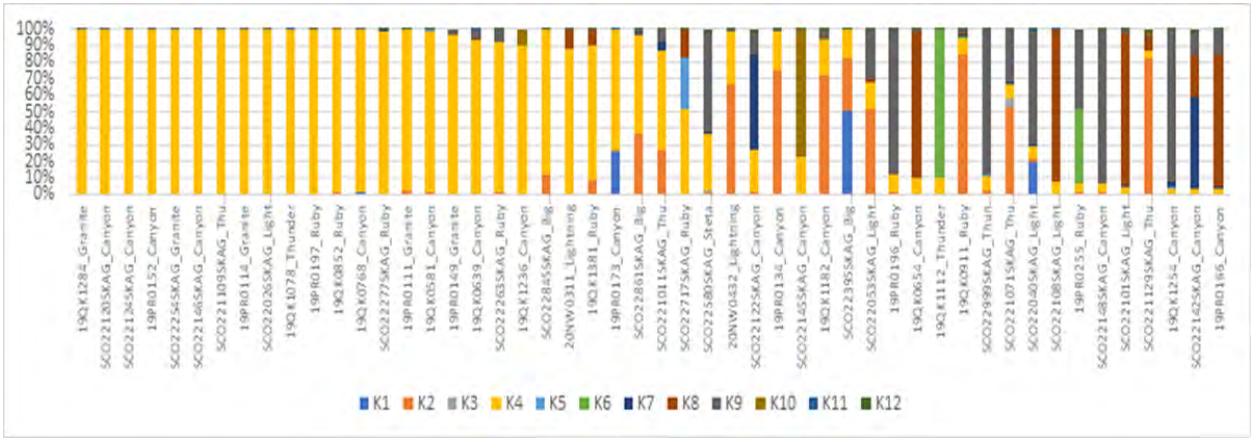


Figure D-5. Posterior sorted by *k*=4.

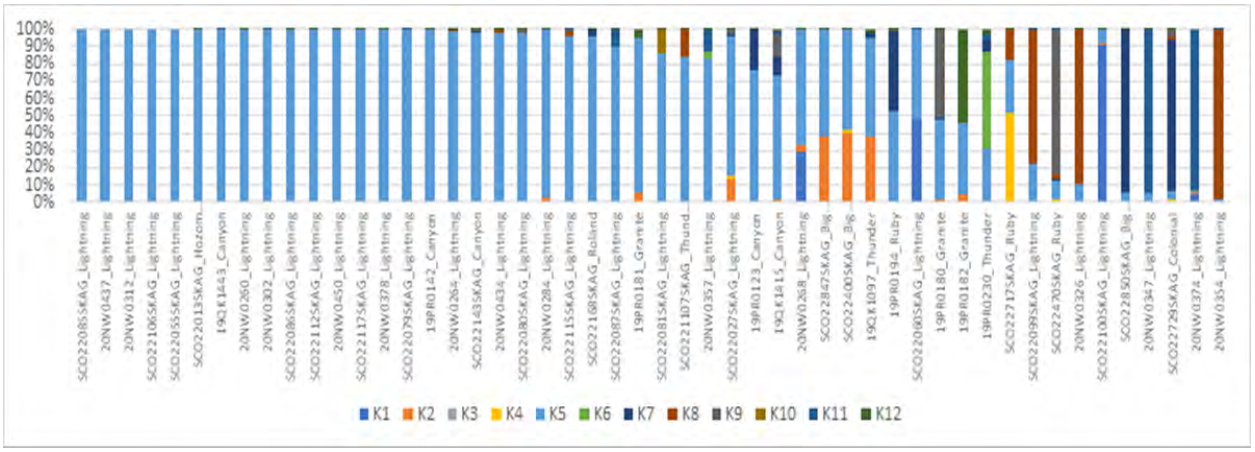


Figure D-6. Posterior sorted by *k*=5.



Figure D-7. Posterior sorted by *k*=6.

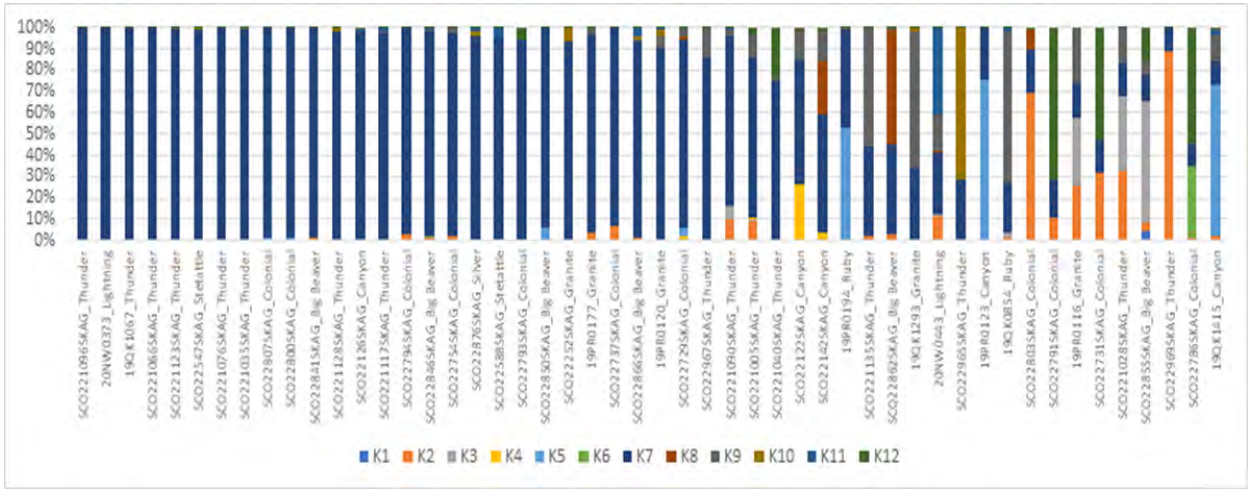


Figure D-8. Posterior sorted by *k*=7.

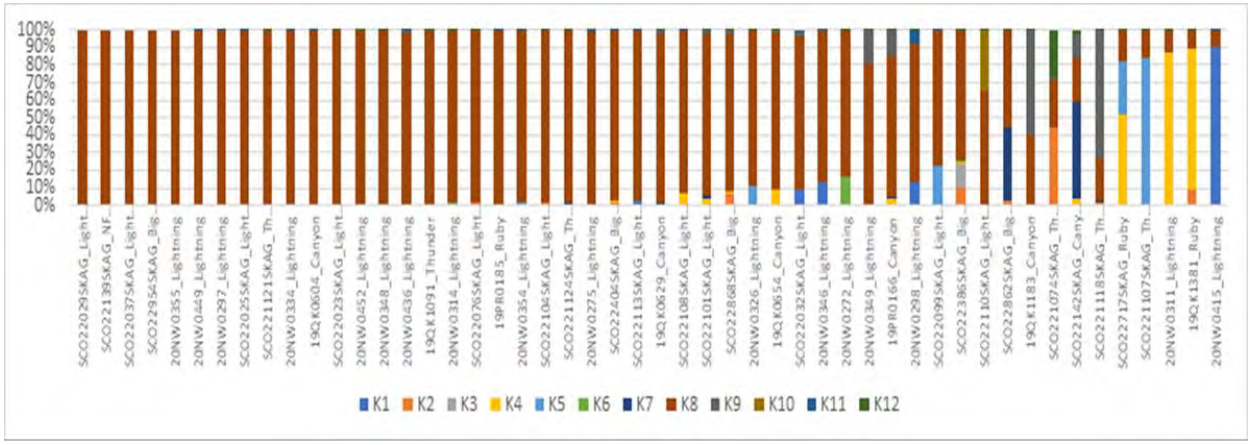


Figure D-9. Posterior sorted by *k*=8.

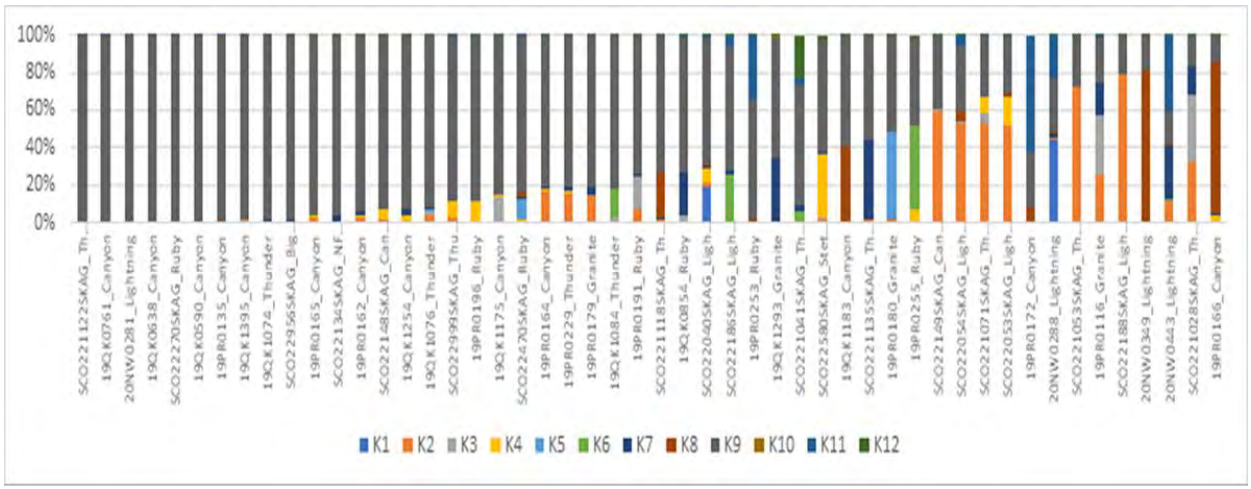


Figure D-10. Posterior sorted by *k*=9.

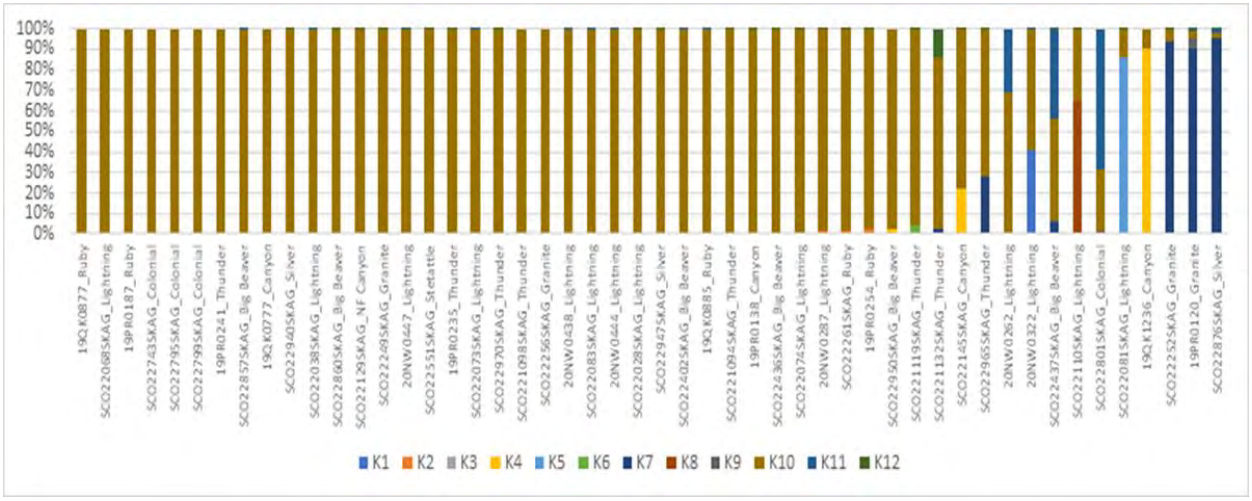


Figure D-11. Posterior sorted by *k*=10.

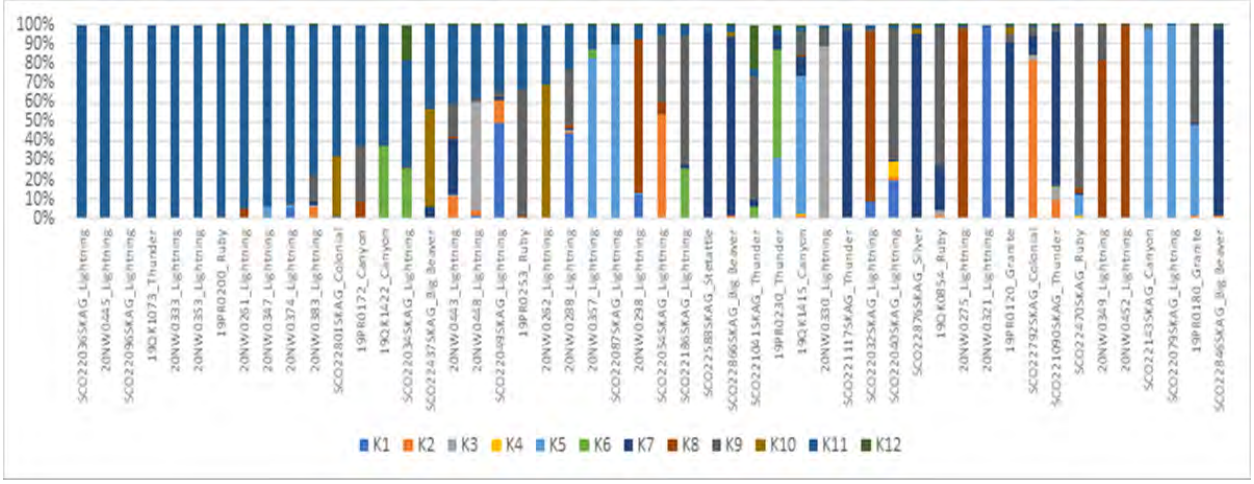


Figure D-12. Posterior sorted by *k*=11.

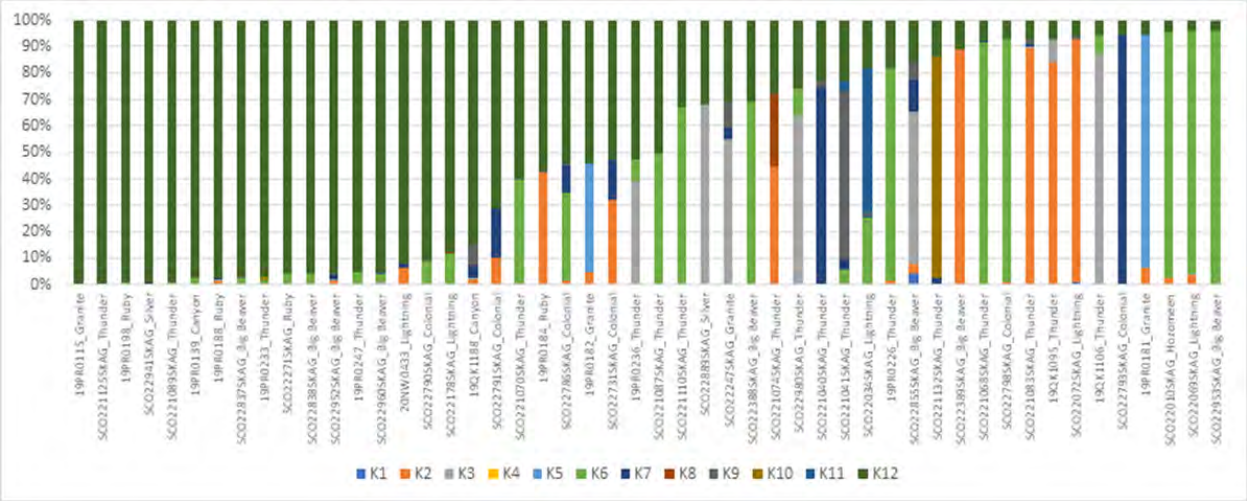


Figure D-13. Posterior sorted by *k*=12.

RESERVOIR NATIVE FISH BASELINE GENETICS STUDY
TECHNICAL MEMORANDUM FOR EXPERT PANEL

ATTACHMENT E

QUANTITATIVE ANALYSIS OF SPECIES SORTED BY K

Table E-1. Count at k_2 for Bull Trout.

	k1	k2
Diablo	0	3
Gorge	11	12
Ross	5	34

Table E-2. Count at k_3 for Bull Trout.

	k1	k2	k3
Diablo	3	0	0
Gorge	10	10	3
Ross	18	3	18

Table E-3. Count at k_4 for Bull Trout.

	k1	k2	k3	k4
Diablo	0	1	0	2
Gorge	10	2	2	9
Ross	2	10	14	13

Table E-4. Count at k_12 for Dolly Varden.

	k1	k2	k3	k4	k5	k6	k7	k8	k9	k10	k11	k12
Big Beaver	0	6	4	3	0	9	5	4	1	5	1	5
Canyon	2	2	1	10	4	2	4	3	13	2	2	2
Colonial	0	4	0	0	0	2	9	0	0	3	1	3
Granite	1	1	2	5	1	0	3	0	4	2	0	3
Hozomeen	0	1	0	0	1	1	0	0	0	0	0	0
Lightning	34	9	4	3	22	4	1	27	4	10	16	2
NF Canyon	0	1	0	0	0	0	0	1	1	1	0	0
Pierce	0	1	0	0	0	0	0	0	0	0	0	0
Roland	0	0	0	0	0	0	1	0	0	0	0	0
Ruby	0	2	1	6	0	1	1	1	6	5	2	4
Silver	0	1	0	0	0	0	1	0	0	2	0	2
Stetattle	1	0	1	0	0	1	2	0	1	1	0	0
Thunder	1	17	12	3	1	18	13	3	9	5	1	9

RESERVOIR NATIVE FISH BASELINE GENETICS STUDY
TECHNICAL MEMORANDUM FOR EXPERT PANEL

ATTACHMENT F

SPECIES IMAGES



Image F-1. Genetically identified Dolly Varden



Image F-2. Genetically Identified Brook Trout



Image F-3. Genetically identified Bull Trout



Image F-4. Genetically identified Bull x Dolly Varden hybrid



Image F-5. Genetically identified Brook x Dolly Varden Hybrid

RESERVOIR FISH GENETICS STUDY

ATTACHMENT D

EXPERT PANEL MEETING SUMMARIES

Seattle City Light **MEETING AGENDA**

Skagit Hydroelectric Project Relicensing Meeting
FA-06 Genetics Expert Panel Meet & Greet
October 26, 2021 1:00 pm – 3:00 pm

Meeting Summary

Disclaimer: These notes are provided to serve as a high-level summary of the meeting and as a communication tool for the benefit of work group continuity. They are streamlined and focused on action items, unresolved issues, future discussion items, and high-level discussion points. They are not intended as a formal record of the meeting.

Attendance

Licensing Participants (LPs):
Alphabetical by last name
Brock Applegate, Washington Department of Fish and Wildlife (WDFW)
Richard Brocksmith, Skagit Watershed Council / Skagit Environmental Endowment Commission (SEEC)
Steve Copps, National Marine Fisheries Services (NMFS)
Jeffrey Garnett, U.S. Fish and Wildlife Service (USFWS)
Brian Lanouette, Upper Skagit Indian Tribe (USIT)
Ashley Rawhouser, National Park Service (NPS)
Kara Symonds, Skagit County
Amy Trainer, Swinomish Tribal Community

Expert Panel:
Hope Draheim, U.S. Fish & Wildlife Service (USFWS)
Jason Dunham, U~~nited~~. S~~tates~~. Geological Survey (USGS)
Alex Fraik, National Marine Fisheries Services (NMFS) Affiliate
Meryl Mims, Virginia Tech
Jim Myers, National Marine Fisheries Services (NMFS)
Krista Nichols, National Marine Fisheries Services (NMFS)
Carl Ostberg, U~~nited~~. S~~tates~~. Geological Survey (USGS)
George Pess, National Marine Fisheries Service (NMFS)
Todd Seamons, Washington Department of Fish and Wildlife (WDFW)
Matt Smith, U.S. Fish & Wildlife Service (USFWS)
Adrian Spidel, NW Indian Fisheries Commission
Rick Taylor, University of British Columbia

Seattle City Light (City Light):
Andrew Bearlin, Seattle City Light
Erin Lowery, Seattle City Light

Consultant Team:
Dan Bingham, Cramer Fish Sciences, Consultant Team
Scott Blankenship. Cramer Fish Sciences~~,~~ Consultant Team
Bao Le, ~~HDR~~ HEC, Consultant Team
Phil Roni, Cramer Fish Sciences, Consultant Team
Erin Settevendemio, HDR, Consultant Team

Facilitation Team:
Greer Maier, Facilitation Team
Lauren Schultz, Facilitation Team

Action Items

Action	Responsibility	Deadline
Licensing Participant (LP) Action Items		
LPs ~~and Expert Panel members~~ will provide existing literature, reports, and analyses used to inform Year 1 of the FA-06 Reservoir Fish Genetics Study to ~~inform the FA-06 baseline fish genetic study~~ LPs and the Expert Panel. These FA-06 Native Fish Genetics resources can be found here.	LPs/~~Expert Panel/~~Triangle	Ongoing
LPs will ~~submit all~~ provide relevant study questions and management objectives~~, suggestions, or concerns~~ for all participants including to the ~~related to the Skagit Relicensing Fish Genetics study to the Facilitation Team. These questions will inform~~ Expert Panel ~~members~~ in advance ~~as they~~ of their review ~~the Tech Memo~~ of the technical memo.	LPs/Triangle	Prior to release of tech memo
City Light/Consultant Action Items		
Dan Bingham (Cramer Fish Sciences) will document the Expert Panel's objectives, function, and overall process.	Dan Bingham/Cramer Fish Sciences, Consulting Team	Complete
Consultant Team will provide existing data, reports, and analyses used to inform Year 1 of the FA-06 Reservoir Fish Genetics Study to LPs and the Expert Panel.	Cramer Fish Sciences	Ongoing

Summary of Issues Discussed, Action Items, and Decisions

Welcome, Introductions, Agenda Overview

The facilitator, Greer Maier, Triangle Associates, and Erin Lowery, Seattle City Light, welcomed the group and led an introduction of attendees. They walked through the agenda and shared the objective of this meeting, which was to allow LPs to meet the Reservoir Native Fish Genetics Study Expert Panel (Expert Panel), learn about their background, and ask questions regarding City Light's study approach and decisions at specific milestones of the Study.

Expert Panel members then briefly described their background and experience ~~in fish genetics~~. The Expert Panel members provide a range of relevant subject matter expertise, including agency and academic geneticists and ecologists.

Expert Panel Members:

- Adrian Spidel, NW Indian Fisheries Commission
- Alex Fraik, NOAA/NMFS Affiliate
- Carl Ostberg, USGS
- George Pess, NOAA/NMFS
- Hope Draheim, USFWS
- Jason Dunham, USGS
- Jim Myers, NOAA/NMFS
- Krista Nichols, NOAA/NMFS

- Matt Smith, USFWS
- Meryl Mims, Virginia Tech
- Rick Taylor, University of British Columbia
- Todd Seamons, WDFW

Study Overview and Expert Panel Role

Erin Lowery, Seattle City Light, provided a brief overview of the FA-06 Reservoir Native Fish Genetics Baseline Study and the Expert Panel goals, milestones, and timelines. The study's goal is to characterize the baseline population genetic structure of the three native fishes present in the reservoir system; bull trout, rainbow trout, and Dolly Varden, and inform long-term planning for native fish management. He outlined specific pieces of the Notice of Agreement that pertain to this study, including:

- ~~A targeted~~ Targeting juvenile fish sampling at spawning grounds for genetics baseline sampling (Year 2).
- Expand~~ed~~ sample collection~~/~~ and/or coordination of existing samples and activities and analysis out-of-basin and above/below dams.
- The Expert Panel role and its expanded field of expertise to include a broad field of experts who are familiar with large landscape processes as well as genetic population structures.

Erin emphasized the critical need for the Expert Panel and their advisory role for future management practices and evaluation of potential effects on fish in the Upper Basin. The purpose of the Expert Panel is to provide input and recommendations to inform City Light's study approach and decisions at specific milestones. The Expert Panel will seek input from LPs to better advise future management decisions with Seattle City Light and fill data gaps related to reservoir fish population genetics. Erin then outlined the Expert Panel milestone timeline, as listed below:

- **Meeting 1 - October 2021**: "Meet and Greet" with City Light and LPs, Q&A.
- **November 2021:** Review of Cramer Fish Sciences Technical Memo detailing available existing data relevant to study.
- **Meeting 2 - December 2021:** Discuss with City Light/LPs the Expert Panel's review of Existing Data Technical Memo, provide recommendations to support the identification of data gaps and Year 2 sampling, and discussion/consideration of additional LP study questions and management objectives.
- **October 2022:** Review draft results of the two-year study.
- **Meeting 3 - November 2022:** Discuss with City Light/LPs the Expert Panel's review of results of the two-year study and provide recommendations on potential topics to be addressed in a long-term reservoir fish and aquatics management plan.

Question and Answer Session with Expert Panel and LPs

The facilitator transitioned the meeting into a question-and-answer session between LPs and the Expert Panel.

- In response to a question regarding the Expert Panel's role in handling disagreements, Erin clarified that the Expert Panel will identify gaps and evaluate research efforts to address those data gaps. If there is a disagreement on what information or analysis is provided, the Expert Panel can outline a path forward.

- Erin clarified that the Expert Panel Scope of Work includes time for review and deliberation on material being presented, and they should be given the latitude to discuss the material as they see fit. The next ~~standing work group meeting with LPs~~ Expert Panel meeting will take place in December and provide the opportunity for LPs and Expert Panel members to discuss the Technical Memo and questions related to the study. The study aims to ensure collaboration with LPs while allowing the Expert Panel to maintain a level of independence as advisors in the process. Outcomes from Expert Panel meetings and deliberations outside of the established timeline will be tracked and shared with LPs. City Light is committed to sharing notes and summarizations of Expert Panel meetings scheduled outside of the established Work Group meetings.
- Dan Bingham, Consultant Team (Cramer Fish Sciences), clarified that the Technical Memo will incorporate Skagit River and Fraser River data and metadata from various studies. The Expert Panel will use this Technical Memo to analyze data gaps that could be addressed by genetic sampling during Year 2 of the study. Erin Lowery noted that a primary goal of the baseline population genetic study is to understand the ~~river system's~~ existing condition~~s and~~ of the native fish populations' genetic structure.
- The Expert Panel requested that the Consultant Team provide as much detail as possible in the Technical Memo and a clear articulation of the expectations of this review exercise. Andrew Bearlin, City Light, explained that the Technical Memo will include all available information and be reviewed in concert with key questions developed by LPs to inform future discussions around specific management practices.
- In response to a comment regarding the gathering and distribution of specific fish genetic questions, the facilitator clarified that the facilitation team will work with LPs to gather questions to provide to ~~the~~ all participants including the Expert Panel ahead of the December meeting.
- Dan Bingham confirmed that any metadata provided in the original data files ~~would be included in the Technical Memo~~ could be provided upon request. The Expert Panel requested information be shared incrementally to inform and advise LPs.
- Jason Dunham, ~~Consultant Team (Cramer Fish Sciences)~~ Expert Panel member (USGS), noted that the Northwest Forest Plan Aquatics Monitoring Report, which includes instream and upslope (roads, forest cover, culverts) responses, may be available next year.

Action Items:

- LPs ~~and Expert Panel members will compile~~ and the consultant team will provide existing ~~literature~~ data, reports, and analyses used to inform Year 1 of the FA-06 Reservoir ~~baseline f~~Fish Ggenetics ~~s~~Study to LPs and the Expert Panel. These FA-06 Native Fish Genetics resources can be found here.
- LPs will provide relevant ~~document all~~ study questions and management objectives~~, suggestions, or concerns into a question inventory spreadsheet~~ for Expert Panel consideration ~~members to contemplate~~ as they begin their review of the technical memo ~~existing fish genetic study materials~~.
- Dan Bingham (Cramer Fish Sciences) will document the Expert Panel's objectives, function, and overall process.

Next Steps

The facilitator reviewed the action items from the meeting and outlined the next steps for the Expert Panel, noting the Expert Panel and LPs can expect to receive the Technical Memo in late November for

review ahead of the next Skagit ~~FA-06 Fish Genetics~~ Expert Panel meeting. The focus of this meeting will be to review and discuss the Reservoir Native Fish Genetics Baseline Study Technical Memo, study questions and management objectives, and data gaps to inform Year 2 sampling ~~and specific LP fish genetic questions~~.

The meeting adjourned at 2:15 pm.

Seattle City Light **MEETING AGENDA**

Skagit Hydroelectric Project Relicensing Meeting
FA-06 Genetics Expert Panel Workshop #2
January 18, 2022

Meeting Summary

Disclaimer: These notes are provided to serve as a high-level summary of the meeting and as a communication tool for the benefit of work group continuity. They are streamlined and focused on action items, unresolved issues, future discussion items, and high-level discussion points. They are not intended as a formal record of the meeting.

Attendance

Alphabetical by last name

Licensing Participants (LPs):

Brock Applegate, Washington Department of Fish and Wildlife (WDFW)
Richard Brocksmith, Skagit Watershed Council / Skagit Environmental Endowment Commission (SEEC)
Pauline Douglas, Nlaka'pamux Nation Tribal Council
Kevin Duncan, Nlaka'pamux Nation Tribal Council
Jeffrey Garnett, U.S. Fish and Wildlife Service (USFWS)
Brian Lanouette, Upper Skagit Indian Tribe (USIT)
Ashley Rawhouser, National Park Service (NPS)
Dudley Reiser, Kleinschmidt Associates/Swinomish Indian Tribal Community
Stan Walsh, Skagit River System Cooperative

FA-06 Expert Panel:

Hope Draheim, ~~U.S. Fish & Wildlife Service (~~USFWS~~)~~
Jason Dunham, U.S. Geological Survey (USGS)
Alex Fraik, National Marine Fisheries Service~~s~~ (NMFS) Affiliate
Meryl Mims, Virginia Tech
Jim Myers, NMFS
Krista Nichols, NMFS
Carl Ostberg, ~~U.S. Geological Survey (~~USGS~~)~~
George Pess, ~~National Marine Fisheries Service (~~NMFS~~)~~
Todd Seamons, ~~Washington Department of Fish and Wildlife (~~WDFW~~)~~
Matt Smith, ~~U.S. Fish & Wildlife Service (~~USFWS~~)~~
Adrian Spidel, NW Indian Fisheries Commission
Rick Taylor, University of British Columbia

Seattle City Light (City Light):

Andrew Bearlin, City Light
Erin Lowery, City Light
Jeff Fisher, City Light

Consultant Team:

Dan Bingham, Cramer Fish Sciences
Scott Blankenship. Cramer Fish Sciences
Danielle Hanson, HDR
Bao Le, HEC
Erin Settevendemio, HDR,
Matt Wiggs, HDR
Jenna Borovansky, HDR

Facilitation Team:

Greer Maier, Facilitation Team
Lauren Schultz, Facilitation Team

Meeting Materials:

- **Meeting Agenda:** Linked Here
- **Meeting Slide Deck:** Linked Here
- **Reservoir Native Fish Genetics Technical Memo** (Linked Here): Summarizes the existing genetics information available and relevant to the *FA-06 Reservoir Native Fish Genetics Baseline Study*.
- **Research Questions and Management Objectives** (Linked Here): Submitted by Licensing Participants from the National Park Service, Upper Skagit Indian Tribe, Washington Department of Fish and Wildlife, U.S. Fish and Wildlife Service, and the National Marine Fisheries Services *(**Note:** These questions should not be interpreted as officially representing the issues, concerns, or positions of the tribe or agencies they come from)*.
- **FA-06 Native Fish Genetics Study Resources:** SharePoint Site

Action Items

Action	Responsibility	Deadline
Licensing Participant (LP) Action Items		
Licensing Participants (LPs) and Expert Panel members will provide additional literature, reports, data, and analyses that may inform the FA-06 Reservoir Native Fish Genetics Study or help address existing LP questions.	LPs/Expert Panel	Ongoing
City Light/Consultant Action Items		
Consultant Team will compile and distribute metadata for data used in FA-06 Reservoir Native Fish Genetics Study.	Consultant Team	Ongoing
Dan Bingham will share table of pairwise FST comparison for O. mykiss with LPs and the Expert Panel.	Dan Bingham, Cramer Fish Sciences (Consultant Team)	~~As soon as possible~~Complete (distributed to LPs on 1/24)
Facilitation Team and Consultant Team will schedule a separate Expert Panel meeting to allow panel members to debrief and discuss approach for engagement in the FA-06 Reservoir Native Fish Genetics Study.	Consultant Team/Facilitation Team	~~As soon as possible~~Complete (meeting occurred on 2/9)

Discussion Topic
Review Pflug (2013) vs Warheit (2014) data sets for application to FA-06 questions and objectives

Summary of Issues Discussed, Action Items, and Decisions

Welcome, Introductions, Agenda Overview

The facilitator, Greer Maier, Triangle Associates welcomed the group and led an introduction of attendees. She walked through the agenda and shared the objectives of this meeting, which were to discuss the existing genetics data for the FA-06 Reservoir Native Fish Genetics Baseline Study and

examine Licensing Participants proposed research questions to evaluate associated data gaps and determine City Light interest in addressing specific questions as part of FA-06 study.

Existing genetics information

Dan Bingham, ~~Consultant Team~~Cramer Fish Sciences, provided an overview of the existing genetics information contained in the FA-06 Technical Memo, starting by outlining objectives for year one and two of the FA-06 Reservoir Native Fish Genetics study. He then laid out the questions addressed by the tech memo and the limitation of post hoc analysis (see slides 2-5). Dan reviewed Rainbow Trout data from the Pflug (2013) data set, describing the sample selection and statistical methods, as well as clarifying how new inferences may be limited by the design of the original study. He noted that genetic structure was apparent in the dataset and that geography affects that structure. Dan did not test any hypotheses related to the causes of genetic structure or effect of hybridization or hatchery influence (see slides 6-13).

Scott Blankenship, Cramer Fish Sciences, provided information on Bull Trout genetics data, outlining population metric collections from the Project Boundary and Project ~~V~~vicinity (see slide 14). He described the approach for removing juvenile samples to avoid bias. He showed the genetic affinity of the collection using genetic principal components and described the major groupings of the data and how differentiated they were (see slide 16).

Dan explained that the purpose of the Technical Memo is to support the Expert Panel, LPs and City Light as they work to answer specific genetic questions that will inform future planning and genetic management decisions. He noted that if questions cannot be answered with existing genetics information found in the Technical Memo, additional data will need to be collected or identified. Erin Settevendemio, HDR, noted that the Technical Memo will be attached to the Initial Study Report as an appendix, and additional information can be considered in year two of the study.

Questions and Discussion:

- In response to a question from an Expert Panel member about the feasibility of the FA-06 study in addressing past activities of the Project, Dan clarified that the goal of the Technical Memo was to provide information on the types of inferences that can be made, and whether they can address existing questions. The Technical Memo is not intended to be used as a baseline of information to answer questions to inform future management decisions.
- In response to a question from an Expert Panel member regarding genotypes in the Pflug (2013) data, Dan explained that the Pflug (2013) data does not include sample IDs, but the study team observed that number of genotypes appeared to be congruent.
- Representatives from the Expert Panel and USIT discussed the Warheit (2014) data set. USIT noted that the Pflug (2013) data had been amended by Warheit, specifically identifying differences in the hatchery introgression. Dan Bingham, Cramer Fish Sciences, clarified that a list of researchers was predetermined and requested for the FA-06 study, and it is not a peer review. *This item was marked as a future discussion topic.*
- In response to a question from an Expert Panel member regarding the possibility of collecting new data, Dan Bingham, Cramer Fish Sciences, explained that new data could be collected if deemed necessary to answer the questions and objective posed by LPs and City Light.
- In response to a question from an Expert Panel member about including the 1970s WDFW data set, Scott explained that the WDFW data set was not included in the tech memo. Erin Lowery,

City Light, clarified that the microsatellite data was used because it was the most universal data set, but the assessment does not need to be limited to only these data. The goal is to identify data gaps and methods that may be more appropriate to address genetic concerns and questions.

- An Expert Panel member suggested a SNPs study may be more appropriate. In response, Scott Blankenship, Cramer Fish Sciences, clarified that SNPs data was not boundary-wide and didn't seem appropriate for population comparison. Andrew Bearlin, City Light, also requested LPs to bring additional information forward, if relevant and appropriate (data must include similar markers). *This was marked as an action item.*
- There was broad discussion and requests from Expert Panel members, USIT, and NPS to include metadata information from collections and to provide clarity regarding where samples were collected. *This was marked as an action item.*
- ~~An~~Matt Smith, Expert Panel, ~~member~~ noted that the purple clusters shown on slide 16 are likely Dolly Varden and Bull Trout; and any outliers are likely hybrids. He noted that it will be important to know the divergence above and below the dam. An Expert Panel member noted that species ~~divergence~~identification data can be found in the Small (2013) and (2016) reports (found here).
- Expert Panel members discussed using other data and information, including SNPs base reports, additional metadata, DNA, and tissue samples in the FA-06 assessment.
- An Expert Panel member requested the table of pairwise Fst for *O. mykiss* collections be shared with the Expert Panel and LPs. *This was marked as an action item.*

Action Items:

- Consultant Team will compile and distribute metadata for data used in FA-06 Reservoir Native Fish Genetics Study.
- ~~Licensing Participants (~~LPs~~)~~ and Expert Panel members will provide additional literature, reports, data, and analyses that may inform the FA-06 Reservoir Native Fish Genetics Study or help address existing LP questions.
- Dan Bingham, Cramer Fish Sciences, will share table of pairwise Fst comparison for *O. mykiss* with LPs and the Expert Panel.

Discussion Topic:

- Review Pflug 2013 vs Warheit 2014 data sets for application to FA-06 questions and objectives.

Background and Questions of interest

Dan Bingham, ~~Consultant Team~~Cramer Fish Sciences, showed the framework through which the study team is analyzing and answering LP and City Light questions and objectives related to the genetic management of native fish. He mentioned that the framework illustrates how population structure, adaptability, abundance, and diversity all ~~indicate~~ inform overall viability (see slide 21). Dan reviewed City Light's objectives for the FA-06 study, as outlined in the Study Plan, and then described the congruence between City Light objectives and LP questions by examining potential metrics to analyze and answer LP questions.

Todd Seam~~a~~ons, Expert Panel, provided historical background of the Skagit, describing historic and recent hatchery activities, which include non-native hatchery Rainbow Trout and hatchery steelhead stocks (both summer and winter). He explained that the non-native hatchery Rainbow Trout are of

Commented [GU1]: I don't recall Erin (or any other member of SCL or the consultant team) clarifying why only microsatellite data were used.

Commented [S(2R1]: This comment was added by Todd Seamons

Commented [S(3]: It should be noted that multiple members of the expert panel said it would be difficult, if not impossible, to generate new microsatellite data. This is partly why the expert panel suggested SNPs may be more appropriate.

California ancestry, summer steelhead from the Washougal River (Skamania stock), and winter steelhead from Chambers Creek stock (South Sound). He noted that releases of Skamania summer steelhead were brief and far in the past and there hasn't been any hatchery winter steelhead released since 2014. Todd went on to explain that the Ross Lake broodstock is an integrated hatchery program using native Ross Lake broodstock and fish are stocked into Gorge and Diablo.

Todd noted that there are coastal ~~c~~Cutthroat ~~t~~Trout in the lower river. The cutthroat discovered in the field in Ross Lake and tributaries have been identified as Westslope Cutthroat Trout, and the assumption is that they are non-native and a potential problem. Ashley Rawhouser, NPS, provided further background information, explaining that ~~e~~Eastern ~~b~~Brook ~~t~~Trout were stocked in the early 1900s and NPS has since then eradicated most of those populations. There is a record of Westslope Cuthroat Trout stocking into Big Beaver Creek. He noted that there is stocking information from a fisheries report in Canada (Triton 2008 report).

Matt Smith, Expert Panel noted that Bull Trout have a high divergence estimate. He described findings from the Ardren paper (linked here), highlighting findings from above the dam. George Pess, Expert Panel, shared information related to the complex geology of the Upper Skagit and its potential importance to genetic population structure. Ashley brought up the work that Jon Reidel conducted as a geologist for NPS and indicated that the headwater capture events between the Fraser, Skagit, and Okanagon had happened on the order of 2 million years ago.[1]

Commented [GU4]: The NPS report was provided to the Expert Panel for context and has been posted in the "Literature" folder on Triangle's SP site.

Questions and Discussion:

- Representatives from the Expert Panel discussed the metrics and their sufficiency in answering the question proposed by LPs. In response, Dan Bingham, Cramer Fish Sciences, suggested using point estimates for Fst to answer LP question #1.
- An Expert Panel member noted the importance of analyzing and understanding historical data to identify and develop long-term monitoring. Todd Seamons, Expert Panel, clarified that pre-dam samples likely do not exist. Samples from the 1980's tend to be the earliest that exist.
- An Expert Panel member suggested further background might be provided through the ongoing studies that are focused on genetics in the Skagit basin.

The group then discussed LP questions and the metrics used to potentially analyze the questions. A summary of the discussion is as follows:

- In response to discussion from the Expert Panel and NPS regarding the specificity of LP questions and the need for higher order questions that genetic data can inform (such as overall fitness of species), Dan Bingham, Cramer Fish Sciences, explained that the viability framework provides the context which City Light is using to frame their objectives. The metrics associated with LP questions should help reinforce understanding and identify data needs. There was agreement on the need for further discussion on the approach to answering and analyzing LP questions.
- There was broad discussion and agreement on the need for the Expert Panel, LPs and City Light to further define long term genetic goals and objectives to advise future management decisions. A member of the Expert Panel noted the importance of understanding the management consequences and long-term application of the study outcomes to management. Andrew Bearlin, City Light, requested further discussion be had to determine LP goals and objectives related to

population management and determine the information needed to meet those goals and objectives.

- In response to a question from the Expert Panel, regarding the time constraints of the FA-06 study, Erin Lowery, City Light, explained that the assessment's goal is to develop a common framework and genetics baseline to address future management decisions. He acknowledged the time constraint of a two-year study but noted the importance of building a solid foundation to understand what can be done within the given timeframe.
- In response to a question from NPS regarding the link to Structured Decision Making, Jason Dunham, USGS, explained there are similarities, and emphasized the need for a measurable outcome to understand how information will be used in the future. A member of the Expert Panel noted that the study should be able to answer the proposed questions with existing data or additional data but was uncertain about the level of confidence that data would be able to provide in the given timeframe.
- A representative from USFWS noted their interest in developing a robust baseline of spawning populations and determining the viability of populations based on their sizes.
- The Expert Panel requested creating space for themselves to debrief and meet separately to discuss the data, approach, questions, and objectives for the FA-06 study. A representative from USIT noted that they would appreciate updates regarding the outcomes from any Expert Panel deliberations outside of the established timeline. City Light is committed to sharing notes and summarizations of Expert Panel meetings scheduled outside of the established Work Group meetings. *This was marked as an action item.*

Action and Discussion Items:

- Facilitation Team and Consultant Team will schedule a separate Expert Panel meeting to allow panel members to debrief and discuss approach for engagement in the FA-06 Reservoir Native Fish Genetics Study.

Wrap Up and Next Steps

The facilitator reviewed the action items and new discussion topics from the meeting and outlined the next steps for the Expert Panel, noting the facilitation team will be reaching out to Expert Panel members to schedule a follow-up meeting. The Expert Panel will continue to review and discuss the Reservoir Native Fish Genetics Baseline Study Technical Memo, LP questions and City Light management objectives, and data gaps to inform Year 2 sampling.

The meeting adjourned at 12:00 pm.

1. In a follow-up conversation after the FA-06 meeting, Jon Riedel, NPS, confirmed that headwater capture has occurred in the Upper Skagit more than once over the last few millions of years especially during the most recent interglacial period. But most notably the Fraser River was flowing through the Skagit Valley due to a blockage near Chilliwack only 11,600 years ago. This is stated in the introduction to a recent report describing geomorphic landforms in the Skagit River written by Jon and colleagues in 2021. In that passage, they state that the Fraser and Okanagan Rivers ran through the Skagit.

Skagit Hydroelectric Project Relicensing Meeting
FA-06 Genetics Expert Panel Work Session
February 9, 2022

Meeting Summary

Disclaimer: These notes are provided to serve as a high-level summary of the meeting and as a communication tool for the benefit of work group continuity. They are streamlined and focused on action items, unresolved issues, future discussion items, and high-level discussion points. They are not intended as a formal record of the meeting.

Attendance

Expert Panel:
Hope Draheim, U.S. Fish & Wildlife Service (USFWS)
Jason Dunham, U. S. Geological Survey (USGS)
Alex Fraik, National Marine Fisheries Services (NMFS) Affiliate
Meryl Mims, Virginia Tech
Jim Myers, NMFS
Carl Ostberg, USGS
George Pess, NMFS
Todd Seamons, Washington Department of Fish and Wildlife (WDFW)
Matt Smith, USFWS
Rick Taylor, University of British Columbia

Facilitation Team:
Greer Maier, Triangle Associates
Lauren Schultz, Triangle Associates

Action	Responsibility	Deadline
LP Action Items		
Triangle will facilitate communication between Expert Panel members and LPs to further refine their questions by (1) defining the importance or relevancy of each question in relation to the larger goal of the FA-06 study (for example, why is the relationship between Skagit and out of basin populations important to know?) and (2) define the spatial significance of each question (for instance, is the issue of genetic assessment of entrainment required at specific sites?)	Triangle/LPs/Expert Panel Members	As soon as possible
City Light Action Items		
Request that City Light provide a distilled version of all available study information to the Expert Panel (including SNP data).	City Light/ Consultant Team	As soon as possible
Request that City Light add additional literature, reports, data, and analyses that may inform the FA-06 Reservoir Native Fish Genetics Study to the Technical Memo or through a separate document (e.g., SNP data, Warheit 2014).	City Light/ Consultant Team	As soon as possible

Request that City Light develop an updated sample collection map that provides comprehensive, simple, and clear delineations between sample locations mentioned in the consultant genetics report. This map should include a distinct legend, clear marker dots, and identification of the dams. The updated map should make it easy to find sample populations from each figure within the consultant genetic report.	City Light/Consultant Team	As soon as possible
Expert Panel		
Matt Smith, Expert Panel, will share a Bull Trout sample collection map with City Light for reference.	Matt Smith, USFWS	As soon as possible

Welcome and Introductions

The facilitator welcomed the group and provided an overview of the agenda. She shared the objective of the meeting, which was to allow the Reservoir Native Fish Genetics Study Expert Panel (Expert Panel), to discuss the data, approach, questions, and objectives for the *FA-06 Reservoir Native Fish Genetics* study. This meeting was requested by the Expert Panel at the January 18th FA-06 Workshop.

Context and Information Sharing

Todd Seamons and Matt Smith provided context from the January 18th FA-06 meeting regarding O. mykiss and Bull Trout data. Todd explained that many of the LP questions and City Light objectives were addressed, or were intended to be addressed, by the existing study reports. Todd explained that City Light's overall interest is in the genetic structure of O. mykiss populations and genetic relationships in the basin, including hatchery fish upstream and downstream of the dam

Todd explained that the Skagit samples have been used in various reports, the main being Pflug (2013). He noted that chapter 8 covers genetic analysis and chapter 10 covers Ken Warheit's attempt to achieve a more precise estimate of identifying hybrids between hatchery steelhead and natural origin steelhead in the lower river. Based on the microsatellite data, there was very little data to precisely identify hybrids between hatchery fish and wild fish.

This Pflug (2013) report was followed by a SNP analysis in the Warheit (2014) report, which was an update to Chapter 10 in Pflug (2013), where SNP genotypes were used to reevaluate the ability to identify hybrids between hatchery steelhead and natural origin steelhead. Todd mentioned that the questions and objectives posed by CL and LPs may have been addressed (vaguely) by both Pflug (2013) and Warheit (2014). The group noted the need to add the Warheit (2014) SNP data as an additional resource for the analysis. *This was added as an action item as a request to City Light.*

Matt Smith provided the group background information on the Smith (2010) report on Bull Trout. He explained that the Smith (2010) report includes a comprehensive sampling of Bull Trout populations in tributaries below the project to the Sauk. He explained that most of the primary samples from above the dams were received from City Light and the collection isn't entirely comprehensive. The report outlines the convergence between above the dam populations and below the dam populations.

Following the Smith (2010) report, Jon Riedel, National Park Service, questioned whether the populations above the dam were accurate, noting the possibility that the Upper Skagit used to flow North into the Fraser River when the Cordilleran ice sheet receded. Matt subsampled Bull trout collections from Fraser and ran a mitochondrial DNA sequence above and below the dam. Results showed only one haplotype, which was not seen anywhere else in the

United States. Matt mentioned he has presentations on the mitochondrial DNA sequencing that he can distribute. The group them moved on to a discussion of available data and how those data apply to License Participant (LP) questions and City Light study objectives.

Discussion of Objectives and Questions

- Rick Taylor mentioned that he needs to review the existing literature again for clarification and to understand whether the existing data is adequate to answer LP questions.
- Matt Smith asked whether the study will be used as a baseline, and if so, questions should be answered with sequence data augmented by additional sample collections. He also mentioned that the existing mitochondrial DNA sequence and the subset of samples provided in Smith (2010) could answer some of the LP questions.
- The group discussed a request for LPs to further refine their questions by (1) defining the importance or relevancy of their questions in relation to the larger goal of the FA-06 study and (2) define the spatial significance of the question. *This request was marked as an action item*.
    - George provided an example from one of the LP questions regarding the relationship of fish in the Skagit, noting that the question did not articulate any rational or context.
- Todd Seamons noted that the consultant team did not provide enough detail at the January 18th FA-06 meeting or in the Technical Memo for Expert Panel members to understand the adequacy of existing information and data to address LP questions and CL objectives.
- George Pess requested more information from City Light on how the Expert Panel can best use their expertise.
- The group requested an updated map that would provide comprehensive, simple, and clear delineations between sample locations. This map should include a distinct legend, clear marker dots, and identification of the dams. Matt Smith mentioned he has a map with specific Bull Trout locations that he can share as an example. *This request and example were marked as action items*.
    - Rick Taylor suggested the Consultant Team might be provided with some specific recommendations for different kinds of analyses by the EP following refinement of questions posed by the LP. Others added that it may be useful for City Light to provide a somewhat distilled version of the available information so the Expert Panel can focus on advising the process without needing to read and sift through so much material. *This request was marked as an action item.*
- George Pess asked if the sample designs were spatially appropriate for the questions being posed. Todd explained that there was thoughtful and strategic planning that went into the sample design and data collection. He mentioned that the execution of the study was not perfect and there were some opportunistic elements to the sample design.
- Todd explained that existing SNP data shows Ross and Gorge Lake Rainbow Trout may be more aligned with Inland Redband or a coastal/inland redband hybrid. FA-06 currently lacks comparison to any Inland Redband. Todd suggested City Light compare the microsatellite data to the Columbia basin Redband microsatellite data since they used the same standardized marker panel.
- Todd reminded the Expert Panel that time is limited on this study, and they only have until the end of 2022 to implement recommendations.
- The group agreed to wait to comment on how to address LP questions until they are further refined and elaborated on.
- Rick volunteered to speak at the February 15th Reservoir Work Group meeting to provide an update on the outcomes and discussion topics from the Expert Panel work session.

Skagit Hydroelectric Project Relicensing
Expert Panel Coordination Meeting
Thursday, March 31, 2022

Draft Meeting Notes

Disclaimer: These notes are provided to serve as a high-level summary of the meeting and as a communication tool for the benefit of work group continuity. They are streamlined and focused on action items, unresolved issues, future discussion items, and high-level discussion points. They are not intended as a formal record of the meeting.

Attendance

Licensing Participants (LPs):
Jeff Garnett, US Fish and Wildlife Service (USFWS)
Brian Lanouette, Upper Skagit Indian Tribe (USIT)
Ashley Rawhouser, National Park Service (NPS)

FA-06 Expert Panel:
Jason Dunham, U.S. Geological Survey
Alexandra Fraik, National Marine Fisheries Services (NMFS)
George Pess, NMFS
Jim Myers, NMFS
Todd Seamons, Washington Department of Fish and Wildlife (WDFW)
Matt Smith, USFWS
Adrian Spidle, NW Indian Fisheries Commission
Rick Taylor, University of British Columbia

Facilitation Team:
Greer Maier, Triangle Associates
Lauren Schultz, Triangle Associates

Meeting Materials

- Study Questions and Objectives submitted by LPs: Linked Here
- Summary from March FA-06 Expert Panel Meeting: Linked Here
- FA-06 Study Resources: Linked Here

Action Items

Action	Responsibility	Deadline
City Light/Consultant Team Action Items		

LPs will provide additional context to their *Genetics Focused Questions Related to the Management of Fish Stocks Associated with Skagit Hydroelectric Project* document. Their revisions will include additional information and delineation on the biological, management, and spatial elements of their questions and objectives.	LPs	ASAP

Welcome and Introductions

Greer Maier, Triangle Associates, introduced herself as the facilitator for the meeting and led a brief round of introductions. The meeting was convened to address an action item from the March FA-06 Expert Panel meeting to coordinate communication between LPs and the Expert Panel to further refine study questions submitted by LPs.

Review of Questions

The group walked through the document, examining the overview section and specific questions in more detail. They identified aspects of the document and questions that needed greater context, detail, and linkage.

Questions and Discussion:

- George Pess, Expert Panel, provided basic questions that the group could investigate:
    1. What is the genetic makeup of *O. mykiss* in and above Ross Lake? Are they similar to inland redband, are they a hybrid, or do they look like downstream *O. mykiss*?
    2. What are the potential changes that can occur if fish passage were to occur given the genetic baseline in and above Ross Lake?
- Rick Taylor, Expert Panel, suggested adding more context to the overview section of the LPs document. He requested LPs provide a direct connection to the specific questions they are proposing, and how they fit into the overall objective
- The group discussed the suggestion of separating questions into **biological aspects** and **methods or management elements**.
- Jason Dunham, Expert Panel, suggested LPs fit and link their questions into a larger umbrella of viability, not just persistence.
- Ashley Rawhouser, NPS, provided additional context to question #5, explaining that the fundamental question is around whether fish migrating up Stetattle Creek are winter Steelhed or summer run Steelhead, and if the difference between the two can be identified based on life history infomration.
- Jim Myer, Expert Panel, noted that in order to understand how operations may have an impact on populations, there is a need to understand what the current population structure is.
- Brian Lanouette, USIT, suggested the group consider how specific migration timings may be hindered by drawdown, and how that may be important for understating central operational changes.
- Rick suggested LPs add and articulate the spatial extent or scale they are interested in for each question, in addition to how and why it's relevant to the overall objective.
- With respect to question #2, Rick explained that out of basin comparisons will be very complicated unless there are many replications available. Brian added that LPs will likely exclude this question.

Action Items:

- LPs will provide additional context to their *Genetics Focused Questions Related to the Management of Fish Stocks Associated with Skagit Hydroelectric Project* document. Their revisions will include additional information and delineation on the biological, management, and special elements of their questions and objectives.

Meeting Wrap-Up and Next Steps

The facilitator reviewed action item and next steps.

The meeting adjourned at 2:30 pm.

DRAFT

Skagit Hydroelectric Project Relicensing Meeting
FA-06 Reservoir Native Genetics Expert Panel Workshop #3
January 30, 2023

DRAFT Meeting Summary

Disclaimer: These notes are provided to serve as a high-level summary of the meeting and as a communication tool for the benefit of work group continuity. They are streamlined and focused on action items, unresolved issues, future discussion items, and high-level discussion points. They are not intended as a formal record of the meeting.

Licensing Participants (LPs):
Brock Applegate, Washington Department of Fish and Wildlife (WDFW)
Richard Brocksmith, Skagit Watershed Council / Skagit Environmental Endowment Commission (SEEC)
Pauline Douglas, Nlaka'pamux Nation Tribal Council
Jeffrey Garnett, U.S. Fish and Wildlife Service (USFWS)
Rick Hartson, Upper Skagit Indian Tribe (USIT)
David Hawkins, USIT
Grant Kirby, Sauk-Suiattle Indian Tribe
Brian Lanouette, USIT
Mike LeMoine, Skagit River System Cooperative (SRSC)
Steve Lewis, Bureau of Indian Affairs (BIA)
Gary Martson, Trout Unlimited
Bill McMillan, Retired Fisheries Biologist
Dave Price, National Marine Fisheries Services (NMFS)
Ashley Rawhouser, National Park Service (NPS)
Dudley Reiser, Kleinschmidt Associates/Swinomish Indian Tribal Community
Chris Smith, USFWS
Stan Walsh, Skagit River System Cooperative

FA-06 Expert Panel:
Hope Draheim, USFWS
Jason Dunham, U.S. Geological Survey (USGS)
Alex Fraik, NMFS Affiliate
Meryl Mims, Virginia Tech
Jim Myers, NMFS
Krista Nichols, NMFS
Todd Seamons, WDFW
Matt Smith, USFWS
Adrian Spidel, NW Indian Fisheries Commission
Rick Taylor, University of British Columbia (UBC)

Absent Expert Panel Members:
Carl Ostberg, USGS
George Pess, NMFS

Seattle City Light (City Light):
Andrew Bearlin, City Light
Erin Lowery, City Light
Jeff Fisher, City Light
Leska Fore, City Light
Matt Love, Cascadia Law
Chris Townsend, City Light
Debra Smith, City Light

Consultant Team:
Simone Barley-Greenfield, HDR
Hans Berge, Cramer Fish Sciences
Dan Bingham, Cramer Fish Sciences
Scott Blankenship, Cramer Fish Sciences
Andy Lara, Cramer Fish Sciences
Bao Le, HEC
Erin Settevendemio, HDR
Matt Wiggs, HDR
Jenna Borovansky, HDR

Facilitation Team:
Betsy Daniels, Facilitation Team
Lauren Schultz, Facilitation Team

Meeting Materials:

- **Year 2 Technical Memo**
- **Year 1 Technical Memo**
- **Meeting Presentation**
- **FA-06 SharePoint Site**
- **FA-06 Native Fish Genetics Study Resources:** SharePoint Site

Action	Responsibility	Deadline
Expert Panel		
Expert Panel Members will review the Year 2 Technical Memo (PDF linked here) and provide any feedback or questions to City Light.	Expert Panel Members	February 10
Triangle Associates		
To support ongoing clarification of study results, Triangle will poll the Expert Panel for their ability to prepare for and attend another Expert Panel discussion in late March, after the Updated Study Report is complete.	Triangle Associates	***March 2023***
Prepare draft meeting summary and send to participating LPs, City Light, and other attendees for review.	Triangle/Consultant Team	As soon as possible

Introduction

This meeting was the final of three Expert Panel meetings for the FA-06 Reservoir Native Fish Genetics Baseline Study. The purpose of the meeting was to review the findings of Year 2 study activities (i.e., Year 2 Technical Memo) and discuss potential management objectives for consideration in future reservoir fish management.

FA-06 Reservoir Native Fish Genetics Baseline Presentation

Dan Bingham, Cramer Fish Sciences, provided a brief overview of the project, Year 1 activities, Year 2 methods, results, conclusions, and discussion questions for *Salvelinus* species (Bull Trout and Dolly Varden) resulting from Year 2 study activities (see slides 1-45). Expert Panel member Rick Taylor, UBC provided alternative interpretation and analyses.

Scott Blankenship, Cramer Fish Sciences, presented the results, conclusions, and discussion questions for *Oncorhynchus mykiss* (Rainbow Trout) resulting from Year 2 study activities (see slides 46-60).

Clarifying questions from LPs and responses from City Light and the Expert Panel:

Sampling and Methods

- License Participant Grant Kirby, Sauk Suiattle Indian Tribe, asked what time of year samples were collected. City Light confirmed that samples were collected during the summer season.
- License Participant G. Kriby asked if the adfluvial life history may have been missed due to the summer sampling timeframe. City Light worked with the Expert Panel to develop the sampling plan and the focus was on juvenile/YOcY/Year 1 life stages as targets. Additional sampling efforts would be needed to target adfluvial migrations, as this wasn't a focus of the Year 2 sampling approach.

DRAFT

- License Participant Richard Brocksmith, Skagit Watershed Council, asked if City Light is interested in using samples from Canada to answer any study questions? City Light welcomes input on additional data or information to bolster the data set, but that would need to be developed for future studies, as the FA-06 study is now complete. Depending on the goals and objectives of interest in the future, Canadian samples may be useful. It was noted by Expert Panel member Rick Taylor, UBC, that samples from Canada could inform the division of Dolly Varden and anchor the suggestion that there was a north/south division.
- License Participant Ashley Rawhouser, NPS, asked if comparisons have been made between fish collected at the Project and fish from the Marblemount Hatchery Ross broodstock program. City Light acquired data from the steelhead program that was discontinued (coastal *O. mykiss*) but not collected from Ross as part of the current Rainbow Trout broodstock program. Expert Panel member Todd Seamons, WDFW, noted WDFW does not have samples from the 2002 collection, as they were shared with Expert Panel member Alex Fraik, NMFS Affiliate, for the SNPs genotypes from the original captive broodstock collection. WDFW has been collecting fish annually (or nearly so) from those fish. Genotyped 2019 broodstock, and others, are available.
- License Participant G. Kirby asked if Rainbow Trout plants were placed into Ross Lake from British Columbia. T. Seamons (Expert Panel) noted that City Light obtained samples from source populations for Rainbow Trout that were planted in the Upper Skagit but were not planted into Ross Lake. Those samples were genotyped.
- R. Taylor (Expert Panel) asked if the anomalous population in Pyramid Creek is from hatchery fish. City Light observed an upstream barrier, and the watershed is located along the freeway; it would be relatively easy for someone to stock fish into Pyramid Creek.
- License Participant Gary Marston, Trout Unlimited, asked if Pyramid Creek samples were directly compared to Skagit River samples below the Project. The samples City Light collected were compared directly to the Project area, above and below dams, and to one hatchery (Marblemount hatchery steelhead).
- License Participant Mike Lemoine, SRSC, asked if distances can be produced by genetic drift over 100 years, and if regional analysis can include neighboring watersheds and the lower Skagit. The Consultant Team clarified that some tests could produce a response within 100 years but were not performed as it was not an objective of the study. However, the distances observed being produced by genetic drift seems unlikely.

Preliminary Results for Salvelinus

- License Participant Brian Lanouette, USIT, asked if low Fst among Bull Trout populations influence low Ne estimates that reflect artificial isolation among the reservoirs. The Consultant Team explained that it is difficult to know given the small sample size and need for sample pooling. At this time, the data is not appropriate for that type of analysis.
- R. Taylor (Expert Panel) noted that the Dolly Varden data was analyzed using microsatellite data and questioned whether there is a plan to use a longer-term approach in the future (i.e., SNP). The Consultant Team noted that further discussion and consideration could be had about how the data can be used in the future and that the data was at a "time zero" time point. Understanding what to do with these data going forward is the primary next step.
- Expert Panel members Meryl Mims, Virginia Tech, and Jason Dunham, USGS, asked whether there is interest in using simulation-based approaches or explicit spatial analyses beyond the descriptive statistics in the technical memo, as it allows exploration of the consequences of decision alternatives. They explained that it's often difficult to obtain enough samples to answer

DRAFT

everything you want to evaluate. While a simulation is not the end all be all, it is a proactive approach that can help inform decisions and understanding of a system. City Light explained that the original intent of the study was to develop a baseline, a solid foundation, and a starting point. The low number of Bull Trout samples was unanticipated, and if the data is to inform management of Bull Trout, additional samples are likely needed.

- Adrian Spidel, NW Indian Fisheries Commission (Expert Panel), noted that prior literature on Bull Trout in the Sauk River showed that effective size in the Sauk went to infinity, and asked if a large Ne would be expected, as the Fst estimate was low among the three reservoirs, which was surprising. The Consultant Team noted that the Sauk River has relatively high effective size and further examination may be warranted.

Preliminary Results for Oncorhynchus

- R. Taylor (Expert Panel) observed that Project *O. mykiss* are intermediate between the interior and coastal zone, which appears to be a natural phenomenon. This may suggest that the Project area is a contact zone whether it is driven by the Project or not. It is an important aspect to prioritize, and since there are numerous samples from the Project area, it is unlikely to be a fluke (beyond Pyramid). The Project is a contact zone between these two subspecies of *O. mykiss* and valuable to preserve. R. Taylor suggested City Light not alter this contact zone for future operations or fish management.
- License Participant A. Rawhouser asked if there are there similar "contact zones" in British Columbia. R. Taylor (Expert Panel) responded that it is relatively common around the post-mountain crest and offered to share further information with LPs.
- License Participant A. Rawhouser asked why Project area fish are not observed in the coastal/downstream cloud (depicted on slide 56). R. Taylor (Expert Panel) explained that overlap between the three clouds exists and at some point, in the past, there may have been some movement and interaction. It was noted by T. Seamons (Expert Panel) that this topic will need further discussion with respect to explanations of the patterns being observed, alternative interpretations, and how it relates to fish passage.
- It was noted by an Expert Panel member Krista Nichols, NMFS, that the interpretation of Omy05 in the *O. mykiss* population needs to be considered carefully regarding whether it confers anadromy or residency. The Consultant Team added that neither the technical memo nor the Updated Study Report (USR) described life history because it was not evaluated, but that the metadata may help inform this topic. The Southwest Fisheries Science Center has worked on diversity for chromosome 5 and patterns were informative related to the propensity for juveniles to migrate, but as the analysis spread north, that relationship broke down. There is a fair bit of unexplored data (e.g., chromosome 28).
- A. Fraik (Expert Panel) asked if there are distinctive patterns within GREB1L/ROCK1 in the *O. mykiss* collections upriver. The Consultant Team clarified that they did not evaluate GREB1L. It was also confirmed that adaptive loci were filtered out from the analyses,
- License Participant David Price, NMFS, asked whether regional analysis was conducted between downstream fish (below Project) and coastal fish. The Consultant Team assumed downstream fish were coastal fish, and while an above/below analysis was conducted, focusing on below-Project fish was not a study objective. T. Seamons (Expert Panel) added it would be beneficial to go through the genotypes that were provided by WDFW with the coastal subspecies from downstream of Gorge Dam (see slide 56).

DRAFT

- The topic of anadromous fish passage issues, i.e., how it might affect *O. mykiss* and char populations, was saved for a future discussion.
- T. Seamons (Expert Panel) asked if there were additional priorities in addition to the Rainbow Trout regional figure. City Light explained that Pyramid Lake SNP samples are important. Another priority is the SNP panel developed for Bull Trout, which has genotypes, and they were supposed to be uploaded to FishGen. T. Seamons provided these genotypes to City Light following the meeting.
- R. Taylor (Expert Panel) observed that *O. mykiss* data seem reasonably straightforward, and that there is a demonstrated structure and regional patterns that should be maintained. He suggested examining diversity through time for Bull Trout and Dolly Varden.

Technical Memo Comments and Requests

Expert Panel members and LPs provided feedback to City Light and the Consultant Team regarding the Year 2 Technical Memo and made suggestions for a review process. A summary of the discussion is as follows:

- *Expert Panel Comments:*
    a. A. Fraik noted that the memo provided a good overview of methods, however, some of the figures in the technical memo were difficult to ascertain where samples came from, which made it challenging to draw conclusions. There was a request that these types of results be presented.
    b. Expert Panel members will provide a review of the technical memo and send back comments to the technical team.
    c. R. Taylor requested clarification on the commenting process. City Light would appreciate any and all comments in writing. The information and results will be consolidated, and the information will be communicated, clarified, and presented in a consistent manner. City Light will be seeking input from co-managers, as well.
    d. Jim Myers, NMFA requested that presumptive populations be displayed in a dendrogram.

- *License Participant Comments:*
    a. A. Rawhouser requested a copy of the "cloud" plot with the Pyramid group in a unique color, such as black.

Note that City Light requested that all comments and requests from the Expert Panel be made in writing and transmitted to Triangle by February 10, 2023.

Process & Expert Panel Scope

The group clarified a process for review of the Year 2 Technical Memo and discussed questions related to the role of the Expert Panel.

Expert Panel Scope:

- Is this meeting the last opportunity for discussion with the Expert Panel? The purpose of the technical memo and this meeting was to present the most relevant findings for consideration by the Expert Panel and further details will be provided in the USR. The intent of Expert Panel was

DRAFT

to provide input to the study team and City Light and identify recommendations and observations to inform the study. The Expert Panel is not being asked to approve the technical memo.

Year 2 Technical Memo Review Process:

- The Expert Panel disagreed with some of the conclusions and inferences in the technical memo and noted that the memo needs more clarity and consensus before next steps and monitoring can be discussed. City Light acknowledged the Expert Panel's concerns and reiterated that it is not the intent of City Light to ask the Expert Panel to approve the study report, and that City Light wants to hear different interpretations and perspectives of the data analyses.
- City Light is open to expanding their scope of participation, for those that have interest and availability, with respect to the study to include additional review, however, City Light needs to consider the timing and sequencing of the Integrated Licensing Process, as the USR is being drafted in parallel and a final document will be filed in early March. Instead of trying to arrive at a consensus, it would be valuable to share various takes on the information, as a next step.
- License Participant David Hawkins, USIT, requested that City Light further the Expert Panel review per the Scope of Work, which was distributed during the meeting.
- The Expert Panel preferred to provide written comments on the Year 2 Technical Memo rather than an additional meeting at this time. To support ongoing clarification of study results, Triangle will poll the Expert Panel for their ability to prepare for and attend another Expert Panel discussion in late March, after the Updated Study Report is complete.

Wrap Up and Next Steps

The facilitation team will be reaching out to Expert Panel members to outline a timeline for review of the Year 2 Technical Memo.

The meeting adjourned at 4:00 p.m. PST.

DRAFT