Site C Clean Energy Project

Fisheries and Aquatic Habitat Monitoring and Follow-up Program

Fish Genetics Study
2022 Status Report for Bull Trout, Arctic Grayling, Rainbow Trout, and Slimy Sculpin

Construction Year 8 (2022)

Armando Geraldes, PhD
University of British Columbia

Eric B. Taylor, PhD
University of British Columbia

May 25, 2023

EXECUTIVE SUMMARY

BC Hydro is currently constructing the Site C Clean Energy Project (the Project) near the town of Fort St. John in northeastern British Columbia which will be the third hydroelectric dam on the Peace River. BC Hydro developed the Site C Fisheries and Aquatic Habitat Monitoring and Follow-up Program (FAHMFP) in accordance with Provincial Environmental Assessment Certificate Condition No. 7 and Federal Decision Statement Condition Nos. 8.4.3 and 8.4.4 for the Project. To date, the Mon-1b, Task 2c (Site C Reservoir Tributaries Fish Population Indexing Survey), Mon-2, Task 2a (Peace River Large Fish Indexing Survey), Mon-2, Task 2b (Peace River Fish Composition and Abundance Survey), the Contingent Fish Capture and Transport Program, and the Temporary Upstream Fish Passage Facility have collected DNA samples from species of game fish, Bull Trout (*Salvelinus confluentus*), Arctic Grayling (*Thymallus arcticus*) and Rainbow Trout (*Oncorhynchus mykiss*), and three small-bodied species of non-game fishes found in the Local Assessment Area (LAA), Slimy Sculpin (*Cottus cognatus*), Longnose Dace (*Rhinichthys cataractae*), and Redside Shiner (*Richardsonius balteatus*). From September 2018 to September 2021 the laboratory of Eric Taylor at the University of British Columbia (UBC) conducted the first phase of the Site C Fish Genetics Study where we: (a) determined levels and patterns of population structure for Bull Trout, Arctic Grayling and Rainbow Trout in the Peace River and its tributaries, (b) developed genotyping assays for genetic monitoring of the system, and (c) deployed those assays for samples collected in the Peace River from 2016 to 2020. That project has now been extended until the end of December 2025 with the following activities planned: Activity 1) population assignment of Bull Trout, Arctic Grayling and

Rainbow Trout samples collected in the Peace River from 2021 to 2024, Activity 2) development and deployment of medium sized genotyping panels (200 to 300 SNPs) for Bull Trout and Rainbow Trout for demographic analyses, and Activity 3) generation of genome-wide sequence data for three small-bodied non-game species for analyses of patterns and levels of population structure in the LAA prior to river diversion. Here we report on the progress of the Site C Fish Genetics Study from January 1, 2022 to December 31, 2022. The results and findings of the previous project can be found in Geraldes and Taylor (2020, 2021, 2022).

For Activity 1, samples of Bull Trout, Arctic Grayling and Rainbow Trout for population assignment were collected in the Peace River in sampling year 2021 and 578 samples were received at UBC where they have been stored and catalogued.

For Bull Trout, 360 samples were collected in the Peace River in 2021, their DNA has been extracted, and they were genotyped at six loci previously developed for population assignment to either of two genetic groups detected in the LAA (Geraldes and Taylor 2020). One genetic group consists of samples that spawn upstream of the Project (UP) in the Halfway River, and the other consists of samples that spawn downstream of the Project (DP) in the Pine River (Geraldes and Taylor 2020). Of the 360 Bull Trout samples collected in 2021 (including 17 sampled from the Temporary Upstream Fish Passage Facility, TUF), only 17 (4.7%) could not be assigned to one of the two groups with more than 95% confidence. The vast majority of samples were assigned to UP (N=335, 93.1% of all samples) and a small number were assigned to DP (N=8, 2.2% of all samples). Of the 17 Bull Trout collected in the TUF in 2021, all were assigned to the UP group.

For Arctic Grayling, 68 samples were collected in the Peace River in 2021, their DNA has been extracted, and they were genotyped at 11 loci previously developed for population assignment (Geraldes and Taylor 2021). Previous work (Geraldes and Taylor 2021) found that four distinct population groups of Arctic Grayling can be identified in the LAA, each one corresponding to a single tributary where they are known to spawn: the Halfway River and the Moberly River (located UP) and the Pine River and the Beatton River (located DP). A total of 67 fish were assigned to the UP group (98.5%), one to the DP group (1.5%), and no fish were unassigned. All 11 fish from the TUF were assigned to UP.

Finally, for Rainbow Trout, 150 samples were collected in the Peace River in 2021, their DNA has been extracted, and they were genotyped at six loci previously developed for population assignment (Geraldes and Taylor 2022). Previous work (Geraldes and Taylor 2022) found that patterns of population structure for Rainbow Trout in the LAA were complex but that two genetic groups, largely corresponding to ancestry from populations spawning UP and ancestry from groups spawning DP (plus hatchery ancestry), could be identified. Of the 150 samples subject to assignment tests in 2022, 69 (46.0%) were assigned to the UP group, 57 (38.0%) to the DP group, and 24 (16.0%) could not be assigned with at least 95% confidence. A single Rainbow Trout was collected in the TUF and it was assigned to the DP group.

An additional 862 samples of those three species, collected in Peace River tributaries in the LAA in 2021, were received at UBC and catalogued. Extraction and quality control of DNA was performed for all the additional 544 samples of Bull Trout and 277 samples of Rainbow Trout collected from Peace River tributaries in the LAA in

2021; these samples will be used for demographic analyses (Activity 2) in subsequent years. All 41 additional samples of Arctic Grayling collected in Peace River tributaries and received at UBC were catalogued and stored but their DNA was not extracted.

For Activity 3, 652 samples were received at UBC and catalogued. Extraction and quality control of DNA was performed for all Slimy Sculpin (N=323), Redside Shiner (N=226), and Longnose Dace (N=103) collected between 2018 and 2020. We used reduced representation genomic DNA sequencing with genotyping-by-sequencing (GBS) to generate sequence data and genetic variant discovery (single nucleotide polymorphisms, SNPs) for 612 samples that passed quality control. So far, data analysis was performed only for Slimy Sculpin. Two distinct genetic groups of Slimy Sculpin were identified in the LAA, one comprising samples from the Moberly River and the other samples from the Peace River. No genetic differentiation was detected between sampling years, nor between sampling sections of the Peace River. Data analysis for Redside Shiner and Longnose Dace will be performed in subsequent years.

We also performed two small ancillary projects aimed at determining the species of a few samples collected in the LAA. For each project we generated genome-wide sequence data and performed phylogenetic and population analysis and were able to determine that i) three samples identified in the field as Slimy Sculpin were instead Prickly Sculpin samples and b) one sample that was identified in the field either as a Brook Trout (*Salvelinus fontinalis*) or a Bull Trout by Brook Trout hybrid was in fact Brook Trout and not a hybrid with Bull Trout.

ACKNOWLEDGMENTS

This work was supported by a contract from BC Hydro. We thank Nich Burnett, Michael McArthur and Brent Mossop for guidance and helpful discussions. Dustin Ford of Golder Associates, Mark LeRuez of Triton Environmental Consultants, and Palmer Environmental Consulting Group are thanked for assistance with sample collection, provision and orientation. Brian Wu, Faith Pan, and Emma Laqua (UBC-Co-Op Program) provided excellent assistance in the laboratory. Much of the laboratory work was supported by the Natural Sciences and Engineering Research Council of Canada grants awarded to EBT (Discovery and Equipment grant programs). The bioinformatics aspects of the project were supported by the UBC Department of Zoology's Computing Unit (ZCU).

LIST OF ACRONYMS AND ABBREVIATIONS

BB	Blind Creek and Burnt River population group of Rainbow Trout
BP	Base pair
CVE	Cross-validation error
DNA	Deoxyribonucleic Acid
DP	Downstream of the Project
FAHMFP	Fisheries and Aquatic Habitat Monitoring and Follow-up Program
F_{ST}	Fixation index is a measure of genetic differentiation owing to population subdivision among localities (S) relative to the total variation in a sample (T)
GBS	Genotyping-by-sequencing
GVCF	Genomic Variant Call Format
HA	Halfway River population group of Rainbow Trout
K	Number of genetic groups in the Admixture analysis
LAA	Local Assessment Area
ML	Moberly River and Lynx Creek population group of Rainbow Trout
PCA	Principal components analysis
PCR	Polymerase chain reaction
QC	Quality control
SNP	Single nucleotide polymorphism
SS	Slimy Sculpin
TUF	Temporary Upstream Fish Passage Facility
UP	Upstream of the Project

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Appendix I. Details of all Bull Trout samples genotyped at six loci using TaqMan assays, their genotypes at those loci and their assignment to the upstream (UP) or downstream (DP) of the project reference groups. File name: "Appendix1_BT_Assignments_AllYears_BCHydroreport2023.xlsx"

Appendix II. Details of all Arctic Grayling samples genotyped at eleven loci using TaqMan™ assays, their genotypes at those loci, their assignment to the upstream (UP) or downstream (DP) of the project reference groups, and to each of Halfway River, Moberly River, Pine River and Beatton River reference groups. File name: "Appendix2_AG_Assignments_AllYears_BCHydroreport2023.xlsx"

Appendix III. Details of all Rainbow Trout samples genotyped at six loci using TaqMan assays, their genotypes at those loci and their assignment to the upstream (UP) or downstream (DP) of the project reference groups. File name: "Appendix3_RB_Assignments_AllYears_BCHydroreport2023.xlsx"

Appendix IVA. Non-game species library sample information and sequencing results. Additional data presented for Slimy Sculpin are genomic coverage and frequency of missing genotypes. File name: "Appendix4AB_Non-Game_species_results.xlsx". Tab name: "Sample Information"

Appendix IVB. Slimy Sculpin PCA and Admixture results for datasets SNP93SS and SNP290SS. File name: "Appendix4AB_Non-Game_species_results.xlsx". Tab name: "PCA_Admixture_Results"

INTRODUCTION

BC Hydro is currently in the eighth year of construction of the Site C Clean Energy Project (the Project) near the town of Fort St. John in northeastern British Columbia (hereafter referred to as the Local Assessment Area, LAA) which will be the third hydroelectric dam on the Peace River. In 2018, BC Hydro and the laboratory of Eric Taylor at the University of British Columbia (UBC), Department of Zoology, entered into a three-year agreement to apply genomic techniques to facilitate aspects of the mitigation and monitoring plan for the LAA. The work covered by that agreement focused on three important recreational sport fishes: Bull Trout (*Salvelinus confluentus*), Arctic Grayling (*Thymallus arcticus*), and Rainbow Trout (*Oncorhynchus mykiss*) that are common in the LAA.

In September 2021 a new four-and-one half year agreement between the lab of Eric Taylor and BC Hydro took effect. The agreement is divided into three activities: (1) to continue the population assignment work for Bull Trout, Arctic Grayling, and Rainbow Trout from 2021 sample years onwards, (2) to develop and deploy medium sized (200 to 300 loci) genomic assays to monitor critical demographic parameters of Bull Trout and Rainbow Trout (e.g., effective population size), and (3) to complete descriptive population genetic structure work for three species of non-game fishes also found in the LAA, Slimy Sculpin (*Cottus cognatus*), Longnose Dace (*Rhinichthys cataractae*), and Redside Shiner (*Richardsonius balteatus*), in support of Mon-15 (Site C Small Fish Translocation Monitoring Program).

These efforts are directly tied to the Site C Fisheries and Aquatic Habitat Monitoring and Follow-up Program (FAHMFP) that BC Hydro developed in accordance

with Provincial Environmental Assessment Certificate, Schedule B, Condition No. 7 and Federal Decision Statement Condition Nos. 8.4.3 and 8.4.4 for the Project. BC Hydro uses multiple lines of evidence to better understand the population structure, migration and movement patterns of these key fish species in the Peace River and its tributaries. Such evidence includes data from otolith and fin ray microchemistry, radio telemetry, fish distribution, and genetics that are being used to test hypotheses developed to answer management questions posed in the FAHMFP.

Purpose and Objectives

The Site C Fish Genetics Study has three main stated activities: (1) to perform population assignment of samples of Bull Trout, Arctic Grayling and Rainbow Trout collected in the mainstem of the Peace River and from the Temporary Upstream Fish Passage Facility (TUF), (2) to develop and deploy medium sized genotyping assays (200 to 300 loci) for genetic monitoring and demographic analysis of Bull Trout and Rainbow Trout in the entire LAA, and (3) to determine levels and patterns of genetic structure of Slimy Sculpin, Redside Shiner and Longnose Dace prior to river diversion.

Geraldes and Taylor (2020, 2021, 2022) reported on the results of the initial project contributing to the FAHMFP. In those reports, the authors summarized genomic work focused on using genotyping-by-sequencing (GBS) across the genomes of Bull Trout, Arctic Grayling, and Rainbow Trout to resolve differences among samples collected from tributaries of the Peace River. For Bull Trout, the Halfway, Moberly and Pine rivers were the focus of study. For Arctic Grayling, the same three rivers plus the Beatton River were the study systems. In Rainbow Trout, samples were examined from the Halfway, Moberly and Pine rivers, a few smaller tributaries of the Peace River

(Farrell, Lynx and Maurice creeks), the Dinosaur Reservoir (created by Peace Canyon Dam located UP), and three hatchery strains known to be used for stocking of fish in the area (Pennask Lake, Blackwater River, and Fraser Valley Domestic).

Geraldes and Taylor (2020, 2021, 2022) revealed strong genetic differences amongst geographic groups that were exploited to develop six (Bull Trout), six (Rainbow Trout), and 11 (Arctic Grayling) TaqMan™ genotyping assays that differentiated samples collected from the mainstem Peace River in terms of whether an individual fish belonged to a spawning population located upstream of the Project (UP, i.e., Halfway River or Moberly River) or downstream of the Project (DP, i.e., Pine River or Beatton River).

Only about 2% of 858 mainstem Peace River samples of Bull Trout could not be assigned to either the UP or DP spawning groups between 2016 and 2020 with more than 95% confidence (overall, about 94% were assigned to UP, about 4% to DP).

Approximately 1.5% of 198 mainstem Peace River samples of Arctic Grayling could not be assigned to either the UP or DP spawning groups with more than 95% confidence over the 2016-2020 period. The vast majority were assigned to UP (91% UP, 7% DP). For Arctic Grayling, population assignment allowed for the assignment of fish to individual tributaries and those results showed that 86% of fish were assigned to the Moberly River (located UP), 7% to the Pine River (located DP), less than 1% to the Halfway River (located UP) and none were assigned to the Beatton River (located DP). Assignment to individual tributaries resulted in a higher percentage (7%) of fish not being assigned with over 95% confidence.

Similar work on Rainbow Trout revealed genetic differences between UP and DP spawning areas, but also showed that fish from some portions of the Pine River had a genetic signature similar to hatchery fish. Furthermore, the data showed much higher levels of admixture between UP and DP genetic groups. This resulted in a higher percentage (17%) of Rainbow Trout samples that were not able to be assigned to UP or DP groups with 95% or higher confidence between 2018 and 2020. Still, the majority of the LAA samples were assigned to UP (57% vs 26% DP). In all three species, there was little variation in UP vs DP assignment among years.

The current report summarizes the work during the first year of the new study to the end of 2022. Specifically, the report summarizes the results of the three main project activities. For Activity 1, Bull Trout, Arctic Grayling, and Rainbow Trout population assignment work for samples collected in the mainstem of the Peace River in 2021 and provides a summary for all sample years between 2016 and 2021. For the demographic analyses within Activity 2, DNA extractions of Bull Trout and Rainbow Trout from all sampling sites in the LAA were completed and progress was made on development of new genomic tools for demographic analyses. This work will be completed in 2023 when it will be deployed to genotype samples collected up to 2021. For Activity 3, we report on the generation of genetic data for analyses of population structure in Slimy Sculpin, Redside Shiner and Longnose Dace and on the analysis of population structure of the samples of Slimy Sculpin collected from the mainstem Peace River and the Moberly River. Reporting of population genetic structure of Redside Shiner and Longnose Dace will follow in 2023. We also report on two ancillary tasks related to

molecular-based identification of several fish sampled in the field whose initial morphology-based identification was uncertain.

BULL TROUT

Materials and Methods

A total of 904 Bull Trout genetic samples were collected from the LAA in 2021 (Table 1). Subsequent DNA extraction and quality control (QC) of all 904 samples followed Geraldes and Taylor (2020). A total of 360 of these samples were used in population assignments (Activity 1); the 544 samples collected in the LAA outside the mainstem of the Peace River (Table 1) were also extracted and will be used in new assays being developed to monitor demographic parameters in populations (Activity 2) that will be described and reported on in 2023.

Geraldes and Taylor (2020) used genome wide polymorphism data generated through GBS to investigate levels and patterns of population structure of Bull Trout in the LAA and determined that there were two population groups in the area, one represented by samples of fish spawning in the Halfway River watershed (located UP) and one by samples of fish spawning in the Pine River watershed (located DP). The authors developed six TaqMan™ assays that allow for the quick and efficient genotyping of six ancestry informative SNPs (i.e., loci showing large levels of genetic differentiation between UP and DP genetic groups) and the assignment of fish to the UP and DP genetic groups. Here, we used those six TaqMan™ assays to genotype the 360 Peace River Bull Trout genetic samples collected in 2021 following the methods in Geraldes and Taylor (2020). Those genotype data were used to assign all Bull Trout samples to spawning tributaries UP and DP following Geraldes and Taylor (2021). Briefly, we used the program GeneClass2 (Piry et al., 2004) to assign samples to UP and DP following the method of Rannala and Mountain (1997). Samples were

considered assigned to UP or DP if they had 95% or higher chance of being from one of those respective groups and considered unassigned if the chance of belonging to either group was lower than 95%.

Table 1. Bull Trout, Arctic Grayling and Rainbow Trout samples available for genetics work for Study year 2021 and across all Study years (2016-2021). Indicated are numbers of samples received (UBC), with DNA extracted (DNA) and genotyped at ancestry informative SNPs (TaqMan).

			Study years 2016-2021			Study year 2021 only		
Species	**Watershed**	**River/SectionID**	**UBC**	**DNA**	**TaqMan**	**UBC**	**DNA**	**TaqMan**
All	All	All	5487	5326	2267	1440	1399	578
Bull Trout	All	All	3514	3514	1288	904	904	360
Bull Trout	Peace River	TUF	17	17	17	17	17	17
Bull Trout	Peace River	Section 1	256	256	256	33	33	33
Bull Trout	Peace River	Section 3	364	364	364	74	74	74
Bull Trout	Peace River	Section 5	319	319	319	177	177	177
Bull Trout	Peace River	Section 6	137	137	137	34	34	34
Bull Trout	Peace River	Section 7	82	82	82	24	24	24
Bull Trout	Peace River	Section 9	31	31	31	1	1	1
Bull Trout	Halfway River	Chowade River	994	994	16	213	213	0
Bull Trout	Halfway River	Colt Creek	28	28	13	10	10	0
Bull Trout	Halfway River	Cypress Creek	849	849	13	200	200	0
Bull Trout	Halfway River	Fiddes Creek	367	367	12	119	119	0
Bull Trout	Halfway River	Halfway River	7	7	6	0	0	0
Bull Trout	Halfway River	Turnoff Creek	40	40	4	0	0	0
Bull Trout	Moberly River	Moberly River	9	9	6	0	0	0
Bull Trout	Peace River	Dry Creek	10	10	10	0	0	0
Bull Trout	Peace River	Maurice	4	4	2	2	2	0
Arctic Grayling	All	All	601	440	311	109	68	68
Arctic Grayling	Peace River	TUF	11	11	11	11	11	11
Arctic Grayling	Peace River	Section 1	4	4	4	0	0	0
Arctic Grayling	Peace River	Section 3	98	98	98	5	5	5
Arctic Grayling	Peace River	Section 5	79	79	79	40	40	40
Arctic Grayling	Peace River	Section 6	42	42	42	6	6	6
Arctic Grayling	Peace River	Section 7	27	27	26	6	6	6
Arctic Grayling	Peace River	Section 9	6	6	6	0	0	0
Arctic Grayling	Beatton River	Beatton River	37	37	3	0	0	0
Arctic Grayling	Beatton River	Bratland Creek	54	53	15	0	0	0

			Study years 2016-2021			Study year 2021 only		
Species	**Watershed**	**River/SectionID**	**UBC**	**DNA**	**TaqMan**	**UBC**	**DNA**	**TaqMan**
Arctic Grayling	Beatton River	La Prise Creek	39	39	13	0	0	0
Arctic Grayling	Beatton River	Unnamed Creek 1	1	1	1	0	0	0
Arctic Grayling	Halfway River	Colt Creek	1	1	1	0	0	0
Arctic Grayling	Halfway River	Kobes Creek	3	0	0	3	0	0
Arctic Grayling	Moberly River	Moberly River	199	42	12	38	0	0
Rainbow Trout	All	All	1372	1372	668	427	427	150
Rainbow Trout	Peace River	TUF	0	0	0	1	1	1
Rainbow Trout	Peace River	Section 1	245	245	245	40	40	40
Rainbow Trout	Peace River	Section 3	243	243	243	61	61	61
Rainbow Trout	Peace River	Section 5	52	52	52	29	29	29
Rainbow Trout	Peace River	Section 6	14	14	14	8	8	8
Rainbow Trout	Peace River	Section 7	21	21	21	11	11	11
Rainbow Trout	Peace River	Section 9	1	1	1	0	0	0
Rainbow Trout	Halfway River	Chowade River	21	21	14	7	7	0
Rainbow Trout	Halfway River	Colt Creek	152	152	12	46	46	0
Rainbow Trout	Halfway River	Cypress Creek	33	33	14	6	6	0
Rainbow Trout	Halfway River	Kobes Creek	243	243	11	93	93	0
Rainbow Trout	Halfway River	Fiddes Creek	1	1	0	1	1	0
Rainbow Trout	Peace River	Dry Creek	7	7	7	0	0	0
Rainbow Trout	Peace River	Farrell Creek	253	253	23	76	76	0
Rainbow Trout	Peace River	Maurice Creek	86	86	11	48	48	0

Results

In 2021, 343 Bull Trout were collected in six sections of the Peace River, and an additional 17 samples were collected from the TUF (Tables 1 and 3; Appendix I). All 360 samples were successfully genotyped at six ancestry informative loci with TaqMan™ assays. As in previous years, most samples were assigned to the UP group (N=335, 93.1%), only eight were assigned to the DP group (2.2% of all samples), and 17 could not be assigned to either group (i.e., assignment probability to either was below 0.95; 4.7% of all samples). Overall, there was little variability in the proportion of fish assigned to UP and DP between 2021 and all previous years (2016 through 2020; Table 2), but

there was a slight increase in the proportion of fish that could not be assigned to either group (4.7% in 2021 versus 2.4% between 2016 and 2020; Table 2). All 17 samples collected from the TUF were assigned to UP.

Table 2. Number of Bull Trout samples collected in the Peace River (PR), including the TUF (Temporary Upstream Fish Passage Facility), and assigned (% of total) to the UP (upstream of the Project) or DP (downstream of the Project) groups with more than 95% confidence based on genotypes at six ancestry informative SNPs.

Location	Year	Total	UP	DP	Unassigned[1]
All Peace River[2]	2021	360	335 (93.1%)	8 (2.2%)	17 (4.7%)
	2016-2020	858	806 (93.9%)	31 (3.6%)	21 (2.4%)
	All years	1218	1141 (93.7%)	39 (3.2%)	38 (3.1%)
TUF	2021	17	17 (100.0%)	0 (0.0%)	0 (0.0%)
PR Section 1	2021	33	32 (97.0%)	1 (3.0%)	0 (0.0%)
	2016-2020	223	216 (96.9%)	4 (1.8%)	3 (1.3%)
	All years	256	248 (96.9%)	5 (1.9%)	3 (1.2%)
PR Section 3	2021	74	72 (97.2%)	1 (1.4%)	1 (1.4%)
	2016-2020	290	266 (91.7%)	10 (3.5%)	14 (4.8%)
	All years	364	338 (92.9%)	11 (3.0%)	15 (4.1%)
PR Section 5	2021	177	159 (89.8%)	5 (2.8%)	13 (7.4%)
	2016-2020	142	131 (92.3%)	9 (6.3%)	2 (1.4%)
	All years	319	290 (90.9%)	14 (4.4%)	15 (4.7%)
PR Section 6	2021	34	32 (94.1%)	1 (3.0%)	1 (2.9%)
	2016-2020	103	94 (91.3%)	8 (7.8%)	1 (1.0%)
	All years	137	126 (92.0%)	9 (6.6%)	2 (1.4%)
PR Section 7	2021	24	22 (91.7%)	0 (0.0%)	2 (8.3%)
	2016-2020	58	58 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	82	80 (97.6%)	0 (0.0%)	2 (2.4%)
PR Section 9	2021	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2020	30	30 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	31	31 (100.0%)	0 (0.0%)	0 (0.0%)

[1]Samples that cannot be assigned to either UP or DP with over 95% confidence

Location	Year	Total	UP	DP	Unassigned[1]

[2]Note that this includes all samples from the LAA analysed this year and in previous years, including some locations not sampled in 2021.

ARCTIC GRAYLING

Materials and Methods

A total of 109 Arctic Grayling samples were collected in 2021 from the LAA (Table 1). Subsequent DNA extraction and QC of all 68 samples collected in the Peace River itself, including 11 from the TUF, followed Geraldes and Taylor (2020).

Geraldes and Taylor (2021) used genome wide polymorphism data generated through GBS to investigate levels and patterns of population structure of Arctic Grayling in the LAA and determined that four population groups could be detected in the area, one for each tributary watershed where Arctic Grayling are known to spawn (Halfway River and Moberly River, both located UP, Pine River and Beatton River, both located DP). The Beatton River and Moberly River population groups were well differentiated from each other and the other localities, but differentiation between the Halfway River and Pine River population groups was less pronounced. The authors developed a set of eleven TaqMan™ genotyping assays targeting ancestry informative SNPs (seven were highly differentiated between UP and DP spawning groups, i.e., Halfway and Moberly versus Pine and Beatton, two were highly differentiated between the Moberly and all others, one was highly differentiated between the Halfway and all others, and one was highly differentiated between the Pine and all others), which resulted in highly successful assignment (98.5%) into UP and DP spawning groups, as well as into each tributary spawning group (92.9%). Here, we used those 11 TaqMan™ assays to genotype the 68 Arctic Grayling samples collected in 2021 from the Peace River, following Geraldes and Taylor (2020). Those genotype data were used to assign Arctic Grayling samples to spawning tributaries UP and DP, as well as to each of the four

spawning tributaries following Geraldes and Taylor (2022). Briefly, we used the program GeneClass2 (Piry et al., 2004) to assign samples to UP and DP following the method of Rannala and Mountain (1997). Samples were considered assigned to UP or DP if they had 95% or higher chance of being from one of those respective groups and considered unassigned if the chance of belonging to either group was lower than 95%.

Results

All samples were successfully genotyped at 11 ancestry informative loci with TaqMan™ assays and could be assigned to the UP (N=67, 98.5%) or DP (one sample collected in sampling Section 6) groups (Tables 1,3, and 4; Appendix II). Though differences between 2021 and previous years were small and likely in part due to differences in sample size (N=68 for 2021 versus N=198 for 2016-2020), the percentage of samples assigned to DP in 2021 was only 1.5%, while it was 7.1% for the previous years (Table 3). All 11 samples collected from the TUF were assigned to UP.

Table 3. Number of Arctic Grayling samples collected in the Peace River (PR), including the TUF (Temporary Upstream Fish Passage Facility), and assigned (% of total) to the UP (upstream of the Project) or DP (downstream of the Project) groups with more than 95% confidence based on genotypes at 11 ancestry informative SNPs.

Location	Year	Total	UP	DP	Unassigned[1]
All Samples	2021	68	67 (98.5%)	1 (1.5%)	0 (0.0%)
	2016-2020	198	181 (91.4%)	14 (7.1%)	3 (1.5%)
	All years	266	248 (93.2%)	15 (5.6%)	3 (1.1%)
TUF	2021	11	11 (100.0%)	0 (0.0%)	0 (0.0%)
PR Section 1	2021	0	0 ()	0 ()	0 ()
	2016-2020	4	3 (75.0%)	1 (25.0%)	0 (0.0%)
	All years	4	3 (75.0%)	1 (25.0%)	0 (0.0%)
PR Section 3	2021	5	5 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2020	93	93 (100.0%)	0 (0.0%)	0 (0.0%)

Location	Year	Total	UP	DP	Unassigned[1]
	All years	98	98 (100.0%)	0 (0.0%)	0 (0.0%)
PR Section 5	2021	40	40 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2020	39	39 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	79	79 (100.0%)	0 (0.0%)	0 (0.0%)
PR Section 6	2021	6	5 (83.3%)	1 (16.7%)	0 (0.0%)
	2016-2020	27	27 (100.0%)	6 (22.2%)	3 (11.1%)
	All years	42	32 (76.2%)	7 (16.7%)	3 (7.1%)
PR Section 7	2021	6	6 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2020	20	16 (80.0%)	4 (20.0%)	0 (0.0%)
	All years	26	22 (84.6%)	4 (15.4%)	0 (0.0%)
PR Section 9	2021	0	0 ()	0 ()	0 ()
	2016-2020	6	3 (50.0%)	3 (50.0%)	0 (0.0%)
	All years	6	3 (50.0%)	3 (50.0%)	0 (0.0%)

[1]Samples that cannot be assigned to either UP or DP with over 95% confidence

As in previous years, when samples are assigned to each of the four spawning tributaries, a larger proportion of samples cannot be assigned with more than 95% confidence to one population compared to assignment as either UP or DP (N=4, 5.9% in 2021 and N=14, 7.1% in previous years, Table 4). Most samples were assigned to the Moberly River population group (91.2%) and only one sample was assigned to either the Halfway or Pine population groups (1.5% to each). As in previous years, no samples were assigned to the Beatton River population group. Ten samples collected from the TUF were assigned to the Moberly River and one was assigned to the Halfway River population group, both located upstream of the Project (Table 4).

Table 4. Number of Arctic Grayling samples collected in the Peace River (PR), including the TUF (Temporary Upstream Fish Passage Facility), and assigned (% of total) to the Halfway River (HA),

Moberly River (MO), Pine River (PI) and Beatton River (BE) with more than 95% confidence based on genotypes at 11 ancestry informative SNPs.

Location	Year	Total	HA	MO	PI	BE	Unassigned[1]
All Samples	2021	68	1 (1.5%)	62 (91.2%)	1 (1.5%)	0 (0.0%)	4 (5.9%)
	2016-2020	198	1 (0.5%)	170 (85.9%)	13 (6.6%)	0 (0.0%)	14 (7.1%)
	All years	266	2 (0.8%)	232 (87.2%)	14 (5.3%)	0 (0.0%)	18 (6.8%)
TUF	2021	11	1 (9.1%)	10 (90.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
PR Section 1	2021	0	0 ()	0 ()	0 ()	0 ()	0 ()
	2016-2020	4	0 (0.0%)	3 (75.0%)	1 (25.0%)	0 (0.0%)	0 (0.0%)
	All years	4	0 (0.0%)	3 (75.0%)	1 (25.0%)	0 (0.0%)	0 (0.0%)
PR Section 3	2021	5	0 (0.0%)	5 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2020	93	1 (1.1%)	86 (92.5%)	0 (0.0%)	0 (0.0%)	6 (6.5%)
	All years	98	1 (1.0%)	91 (92.9%)	0 (0.0%)	0 (0.0%)	6 (6.1%)
PR Section 5	2021	40	0 (0.0%)	38 (95.0%)	0 (0.0%)	0 (0.0%)	2 (5.0%)
	2016-2020	39	0 (0.0%)	38 (97.4%)	0 (0.0%)	0 (0.0%)	1 (2.6%)
	All years	79	0 (0.0%)	76 (96.2%)	0 (0.0%)	0 (0.0%)	3 (3.8%)
PR Section 6	2021	6	0 (0.0%)	5 (83.3%)	1 (16.7%)	0 (0.0%)	0 (0.0%)
	2016-2020	36	0 (0.0%)	26 (72.2%)	5 (13.9%)	0 (0.0%)	5 (13.9%)
	All years	42	0 (0.0%)	31 (73.8%)	6 (14.3%)	0 (0.0%)	5 (11.9%)
PR Section 7	2021	6	0 (0.0%)	4 (66.7%)	0 (0.0%)	0 (0.0%)	2 (33.3%)
	2016-2020	20	0 (0.0%)	15 (75.0%)	4 (20.0%)	0 (0.0%)	1 (5.0%)
	All years	26	0 (0.0%)	19 (73.1%)	4 (15.4%)	0 (0.0%)	3 (11.5%)
PR Section 9	2021	0	0 ()	0 ()	0 ()	0 ()	0 ()
	2016-2020	6	0 (0.0%)	2 (33.3%)	3 (50.0%)	0 (0.0%)	1 (16.7%)
	All years	6	0 (0.0%)	2 (33.3%)	3 (50.0%)	0 (0.0%)	1 (16.7%)

[1]Samples that cannot be assigned to any single population with over 95% confidence.

RAINBOW TROUT

Materials and Methods

A total of 427 Rainbow Trout genetic samples were collected in 2021 from the LAA (Table 1). Subsequent DNA extraction and QC of all 427 samples followed Geraldes and Taylor (2020). A total of 150 of these samples were used in population assignments (Activity 1); the 277 samples collected in the LAA outside the mainstem of the Peace River (Table 1) were also extracted and will be used in new assays being developed to monitor demographic parameters in populations (Activity 2) that will be described and reported on in 2023.

Geraldes and Taylor (2021) generated GBS sequence data to determine levels and patterns of population structure of Rainbow Trout in the LAA. Population genetic analysis of those data (Geraldes and Taylor 2022) revealed a pattern of population structure where three groups were identified, largely corresponding to i) samples collected in the Halfway River (HA), ii) samples collected in the Moberly River and Lynx Creek (ML), and iii) samples collected from tributaries of the Pine River, Blind Creek and Burnt River (BB). The authors noted that the results suggested that there were much higher levels of admixture between these groups than observed for Bull Trout or Arctic Grayling. In particular, all samples from the Pine River proper and Willow Creek (a Pine River tributary), Maurice Creek and Farrell Creek (Peace River tributaries) appeared as complex mixes of the three groups (but predominantly from the ML and BB groups). Some samples collected in the Halfway River were also highly admixed between the three groups. Finally, samples from three hatchery strains, commonly used for stocking, had a genetic signature similar to that of the BB group and this genetic

group contributed to much of the admixture found in fish from all other localities. Geraldes and Taylor (2022) suggested that there may have been some introgression of hatchery strains into Rainbow Trout in the LAA associated with stocking activities in the past. Geraldes and Taylor (2022) developed six TaqMan™ assays that allow for the quick and efficient genotyping of six ancestry informative SNPs, i.e., loci showing large levels of genetic differentiation between UP (HA and ML) and DP (BB) genetic groups, and the assignment of fish to the UP and DP genetic groups. Here, we used those six TaqMan™ assays to genotype the 150 Rainbow Trout genetic samples collected in 2021 from the Peace River, following the methods in Geraldes and Taylor (2022). Those genotype data were used to assign all Rainbow Trout samples to spawning tributaries UP and DP following Geraldes and Taylor (2022). Briefly, we used the program GeneClass2 (Piry et al., 2004) to assign samples to UP and DP following the method of Rannala and Mountain (1997). Samples were considered assigned to UP or DP if they had 95% or higher chance of being from one of those respective groups and considered unassigned if the chance of belonging to either group was lower than 95%.

Results

In 2021, 149 Rainbow Trout were collected in five sections of the Peace River. One sample was collected from the TUF (Tables 1 and 5; Appendix III). All 150 samples were successfully genotyped at six ancestry informative loci with TaqMan™ assays. More samples were assigned to the UP group (N=69, 46% of all samples) than to the DP group (N=57, 38% of all samples). The proportion of samples that could not be assigned to either group (17.1%) was similar to the proportion found for previous years, but the proportion of samples assigned to the DP group was higher than in previous

years (38% for 2021 versus 25.5% for previous years). The one sample collected from the TUF was assigned to DP (Table 5).

Table 5. Number of Rainbow Trout samples collected in the Peace River (PR), including the TUF (Temporary Upstream Fish Passage Facility), and assigned (% of total) to the UP (upstream of the Project) or DP (downstream of the Project) groups with more than 95% confidence based on genotypes at six ancestry informative SNPs.

Location	Year	Total	UP	DP	Unassigned[1]
All Peace River	2021	150	69 (46.0%)	57 (38.0%)	24 (16.0%)
	2018-2020	427	245 (57.4%)	109 (25.5%)	73 (17.1%)
	All years	577	314 (54.4%)	166 (28.8%)	97 (16.8%)
TUF	2021	1	0 (0.0%)	1 (100.0%)	0 (0.0%)
PR Section 1	2021	40	19 (47.5%)	13 (32.5%)	8 (20.0%)
	2018-2020	205	118 (57.6%)	45 (22.0%)	42 (20.5%)
	All years	245	137 (55.9%)	58 (23.7%)	50 (20.4%)
PR Section 3	2021	61	33 (54.1%)	19 (31.1%)	9 (14.8%)
	2018-2020	182	118 (64.8%)	37 (20.3%)	27 (14.8%)
	All years	243	151 (62.1%)	56 (23.0%)	36 (14.8%)
PR Section 5	2021	29	16 (55.2%)	7 (24.1%)	6 (20.7%)
	2018-2020	23	8 (34.8%)	11 (47.8%)	4 (17.4%)
	All years	52	24 (46.2%)	18 (34.6%)	10 (19.2%)
PR Section 6	2021	8	0 (0.0%)	8 (100.0%)	0 (0.0%)
	2018-2020	6	0 (0.0%)	6 (100.0%)	0 (0.0%)
	All years	14	0 (0.0%)	14 (100.0%)	0 (0.0%)
PR Section 7	2021	11	1 (9.1%)	9 (81.8%)	1 (9.1%)
	2018-2020	10	1 (10.0%)	9 (90.0%)	0 (0.0%)
	All years	21	2 (9.5%)	18 (85.7%)	1 (4.8%)
PR Section 9	2021	0	0 ()	0 ()	0 ()
	2018-2020	1	0 (0.0%)	1 (100.0%)	0 (0.0%)
	All years	1	0 (0.0%)	1 (100.0%)	0 (0.0%)

[1]Samples that cannot be assigned to either UP or DP with over 95% confidence.

NON-GAME FISH SPECIES

Materials and Methods

Samples

A total of 652 tissue samples of three non-game fish species present in the LAA (Table 6; Appendix IV) were collected up to 2020, i.e., prior to river diversion: 323 Slimy Sculpin (*Cottus cognatus*), 226 Redside Shiner (*Richardsonius balteatus*), and 103 Longnose Dace (*Rhinichthys cataractae*). Each tissue sample was stored in an individual vial with 95% ethanol and shipped to UBC for genetic analysis. Genetic analyses of these samples support Activity 3 of the agreement with BC Hydro. The procedures for DNA extraction and QC followed Geraldes and Taylor (2020).

Table 6. Number of samples of Slimy Sculpin (SS), Redside Shiner (RS) and, Longnose Dace (LD) collected in the LAA for which DNA was extracted (UBC), number of samples used for sequencing (GBS), and number of samples used in population genetic analysis (SNP293 and SNP290).

Species	River/SectionID	Year	UBC[1]	GBS[2]	SNP293[3]	SNP290[4]
All	All	All	652	612	NA	NA
SS	All	All	323	299	293	290
SS	Peace River-Section 3	2018	20	20	20	20
SS	Peace River-Section 3	2019	21	21	20	20
SS	Peace River-Section 3	2020	15	15	15	15
SS	Peace River-Section 5	2018	19	19	18	16
SS	Peace River-Section 5	2019	20	20	19	18
SS	Peace River-Section 5	2020	89	88	89	89
SS	Peace River-Section 7	2020	80	80	78	78
SS	Moberly River	2018	11	11	11	11
SS	Moberly River	2019	23	1	1	1
SS	Moberly River	2020	24	23	22	22
SS	Halfway River-Cypress Creek	2018	1	1	0	0
RS	All	All	226	218	NA	NA
RS	Peace River-Section 3	2018	20	20	NA	NA
RS	Peace River-Section 3	2019	20	20	NA	NA

Species	River/SectionID	Year	UBC[1]	GBS[2]	SNP293[3]	SNP290[4]
RS	Peace River-Section 3	2020	25	25	NA	NA
RS	Peace River-Section 5	2018	23	23	NA	NA
RS	Peace River-Section 5	2019	20	20	NA	NA
RS	Peace River-Section 5	2020	33	33	NA	NA
RS	Peace River-Section 7	2020	4	4	NA	NA
RS	Moberly River	2018	20	20	NA	NA
RS	Moberly River	2019	23	15	NA	NA
RS	Moberly River	2020	27	27	NA	NA
RS	Peace River-Lynx Creek	2016	11	11	NA	NA
LD	All	All	103	95	NA	NA
LD	Peace River-Section 3	2019	3	3	NA	NA
LD	Peace River-Section 5	2019	5	5	NA	NA
LD	Peace River-Section 5	2020	7	7	NA	NA
LD	Peace River-Section 7	2020	8	8	NA	NA
LD	Moberly River	2018	20	20	NA	NA
LD	Moberly River	2019	24	16	NA	NA
LD	Moberly River	2020	34	34	NA	NA
LD	Peace River-Maurice Creek	2006	2	2	NA	NA

[1]Number of samples received at UBC

[2]Number of samples used for genotyping-by-sequencing (GBS)

[3]Number of samples used for population genetic analysis.

[4]Number of samples used for population genetic analysis after eliminating three samples that were identified by genetic tools as Prickly Sculpin (see text for details).

Sequencing

We used reduced representation genomic DNA sequencing with genotyping-by-sequencing (GBS) to generate sequence data and genetic variant discovery (single nucleotide polymorphisms, SNPs). We largely followed the protocol previously used for Bull Trout (Geraldes and Taylor 2020), and Arctic Grayling and Rainbow Trout (Geraldes and Taylor 2021). Here, we detail sample selection for each species and indicate additional changes made to the protocol to sequence in a single Illumina NovaSeq 6000 S4 PE150 run with 672 samples and not the 96 samples ran previously

for Bull Trout, Arctic Grayling and Rainbow Trout. This change in approach was prompted by increasing read throughput obtained with Ilumina NovaSeq 6000 S4 PE150 (approximately 3 billion paired end reads versus the approximately 200-400 million paired end reads obtained with Illumina HiSeq4000 PE150 used for Bull Trout and Arctic Grayling or the approximately 500 million reads obtained with Illumina NovaSeq 6000 SP PE150 used for Rainbow Trout sequencing).

All samples of each of the three non-game fish species for which DNA was successfully extracted and passed QC (293 Slimy Sculpin, 218 Redside Shiner, and 95 Longnose Dace; Table 6 and Results section) were used for GBS. We also included one sample of Coastrange Sculpin (*Cottus aleuticus*), a closely related species to Slimy Sculpin, one sample of Longnose Dace from an allopatric population in BC, one sample of Redside Shiner from an allopatric population in Oregon, and one sample of Nooksack Dace, a divergent form of Longnose Dace. These additional samples were included to provide some taxonomic or geographic scale to help interpret the levels of differentiation observed within the LAA. In total we prepared GBS sequencing libraries for 658 samples and 14 negative controls distributed over 7 plates with 96 wells each.

While in our previous GBS library preparation protocol, samples were barcoded with 96 individual sample barcodes (a unique 4 to 8 bp combination of nucleotides), for our new protocol we designed 8 additional 5 to 8 base pair plate barcodes which, when used in combination with the sample barcodes, allowed us to multiplex, pool, and sequence in one single reaction up to 768 samples.

Specifically, for each sample, we digested 100 ng of genomic DNA with the enzyme *Pst*I (New England Biolabs, Ipswhich, MA, USA) at 37°C for 3 hours in the

presence of both the sample and plate barcodes. Next, to attach the barcodes to the digested DNA fragments, all three components were ligated with T4 DNA ligase (New England Biolabs, Ipswhich, MA, USA) following the manufacturer's instructions for 1 hour at 22°C followed by enzyme inactivation at 65°C for 10 minutes. The resulting reactions were then cleaned with AMPure XP beads (Beckman-Coulter, Brea, CA, USA) with a 30:20 beads:ligated DNA solution to remove DNA fragments smaller than 100 bp (including excess non-ligated sample and plate barcodes) as well as other chemicals in the solution that inhibit the subsequent PCR reaction. Purified DNA was eluted in 25 μL of AE buffer (Qiagen Inc., Valencia, CA, USA) of which 6 μL were used for PCR amplification with Phusion High-Fidelity DNA polymerase (New England Biolabs, Ipswhich, MA, USA) at a final reaction volume of 25 μL. Individually barcoded samples were amplified via PCR (PCR mix followed the manufacturer's instructions) with the following program consisting of 18 amplification cycles of 98°C for 10 s, 65°C for 30 s and 72°C for 40 s, preceded by an initial DNA denaturation for 30 s at 98°C and followed by a final DNA extension for 5 m at 72°C. We ran 2 μL of each PCR amplified DNA to check for a DNA smear indicating that there was no preferential amplification of some fragment sizes but rather that a large range of product sizes were amplified. The amplified DNA was quantified with Qubit dsDNA high sensitivity kit (Thermo Fisher Scientific, Waltham, MA, USA) and 250 ng of each sample's barcoded DNA was added to a common pool (final volume of the DNA pool was approximately 6.8 mL). The DNA pool was then cleaned and concentrated by adding 7.5 mL of AMPure XP beads (Beckman-Coulter, Brea, CA, USA) following the manufacturer's protocol. The DNA pool was eluted in 40 μL of AE buffer (Qiagen Inc., Valencia, CA, USA). We ran the

concentrated pooled library over several lanes (3 μg of DNA pool per lane) of a 2% agarose gel stained with 1% SYBR safe DNA gel stain (Thermo Fisher Scientific, Waltham, MA, USA) at 90V for 2 hours and then excised the 500-700 bp gel section from each lane. The DNA was extracted and purified from the agarose gel with the QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA) and the size distribution of the fragments in the library was checked in an Agilent High Sensitivity DNA chip ran on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The pooled library was sequenced in an Illumina NovaSeq 6000 S4 with 150 bp paired end reads at the McGill University and Génome Québec Innovation Centre.

For Slimy Sculpin we followed a bioinformatics pipeline for GBS read processing, mapping, and variant calling and evaluation, broadly similar to the one used for Bull Trout in this Project (Geraldes and Taylor 2020), which is available at https://doi.org/10.5061/dryad.t951d (Irwin et al. 2016), as well as a few modifications to the pipeline that were already implemented in the analyses of the Arctic Grayling (Geraldes and Taylor 2021) and the Rainbow Trout (Geraldes and Taylor 2022) datasets. For the pooled DNA libraries sequenced in 2022, which used dual barcoding, reads were again demultiplexed with the function “process_radtags” from the STACKS v2.5 pipeline (Catchen et al. 2013) but here, each individual is identified with two barcodes, the sample barcode and the plate barcode, one in each read of a paired end read set.

After demultiplexing, we restricted our analysis to sequences from Slimy Sculpin (299 samples) and Coastrange Sculpin (one sample) for which a reference genome sequence from a congeneric species exists (Dennenmoser et al. 2019). Sequence data

for Longnose Dace and Redside Shiner, for which no genome reference sequences exist, nor for a closely related species, will be analyzed in 2023, with a genome reference free approach (e.g., Catchen et al. 2013).

Sequence reads from Slimy Sculpin and Coastrange Sculpin were aligned to the genome reference sequence (Dennenmoser et al. 2019) of the Rhine sculpin, *Cottus rhenanus,* available online at Dryad (https://doi.org/10.5061/dryad.48g9f5r). The Rhine Sculpin is native to western Europe and is the only *Cottus* for which a reference genome is available. This is a highly contiguous genome reference sequence consisting of 24 chromosome level contigs which encompass 94.8% of the entire reference sequence, and 6,834 smaller contigs that have yet to be anchored to the 24 chromosomes (Dennenmoser et al. 2019). Read trimming, mapping to the reference genome, polymorphism identification and SNP calling followed the protocols successfully employed for the Bull Trout (Geraldes and Taylor 2020) dataset, with modifications previously reported for the Arctic Grayling (Geraldes and Taylor 2021) and Rainbow Trout (Geraldes and Taylor 2022) datasets.

Analyses of Population Structure in Slimy Sculpin

After polymorphism identification, we used VCFtools v0.1.11 (Danecek et al. 2011) to determine mean read depth per sample and levels of missing data across all identified polymorphisms. Seven samples (out of 300) were eliminated from further analysis due to low read numbers, low mean read depth after mapping to the reference genome, and high levels of missing data at putatively polymorphic loci (see Results). One of those samples was the Coastrange Sculpin, the other six were Slimy Sculpin samples from Cypress Creek, the Moberly River and the Peace River mainstem (Table 6).

For the remaining 293 samples (SNP293 dataset, Table 6), we first used a custom script (Owens et al. 2016) to eliminate variants that showed an observed heterozygosity of 0.6 or higher across all retained samples, as these are likely the result of mapping to paralogous regions of the genome and then, using VCFtools v0.1.11 (Danecek et al. 2011), we filtered our polymorphism file further to arrive at a set of high-quality SNPs to form the basis of subsequent population genetic analysis. Namely, we eliminated: i) insertion/deletion polymorphisms to retain only SNPs, ii) SNPs with more than two alleles, iii) SNPs with genotype quality below 10 (these have a higher than 10% chance of being incorrect genotypes), iv) SNPs with missing genotypes in more than 30% of samples, and v) low frequency SNPs (SNPs present at a frequency below 1%). For analysis of population structure (see below), we used Plinkv1.9 (Chang et al. 2019) to remove SNPs that were in close linkage with other SNPs in the set (option "--indep-pairwise 50 10 0.2" to eliminate SNPs with r^2 greater than 0.2 in overlapping windows of 50 consecutive SNPs moving 10 SNPs at a time between windows) as they are not independent data points.

Analysis of population structure, see below in the Results section, suggested that three samples (all from Section 5 of the Peace River) might have been misidentified in the field and belong to a different species. We generated a second SNP dataset by removing those three samples (after confirming their mis-identification – see below) and repeated filtering described above for the remaining 290 samples (SNP290 dataset, Table 6).

Following the analysis pipeline previously employed for Bull Trout (Geraldes and Taylor 2020), Arctic Grayling (Geraldes and Taylor 2021), and Rainbow Trout (Geraldes

and Taylor 2022) we used two complementary and independent approaches to infer patterns of population structure in Slimy Sculpin. In the first approach, we ordinated the SNP dataset in “genotype space” using principal components analyses (PCA) with the R package SNPrelate (Zheng et al. 2012) to summarize genetic variation into up to ten successive orthogonal principal components (PCs). In the second approach, we used the program Admixture v1.3.0 (Alexander et al. 2009) to estimate ancestry proportions for each fish. Admixture is a program that models the probability of the observed genotypes using ancestry proportions and population allele frequencies with a maximum likelihood approach to determine the most likely number of genetic groups (i.e., K). In this analysis, individual fish can be composed of more than one of these K genetic groups and the analysis provides an estimate of the proportion of each fish’s genome composed of each of the K groups (i.e., its admixture proportions). To assess the consistency of the results we ran five replicates of Admixture for each K from one to seven and terminated each run when the difference in log-likelihood between successive iterations fell below 1×10^{-9}. We chose the K value that minimized the cross-validation error (CVE), i.e., that best fit the data (Alexander et al. 2009), and made one last run with K varying from two to four using 1,000 bootstraps to estimate the standard error of the inferred admixture proportions for each K.

We used VCFtools (Danecek et al. 2011) to estimate per locus Weir and Cockerham’s F_{ST} (Weir and Cockerham, 1984) to quantify levels of genetic differentiation between sampling regions and between sampling years. These analyses were performed for the SNP290 Slimy Sculpin dataset for all high-quality SNPs after eliminating SNPs with minor allele frequency below 1%, but prior to eliminating SNPs in

close linkage disequilibrium (N= 38,265 SNPs in total were used to generate F_{ST} estimates).

Results

We successfully extracted large enough amounts of DNA with sufficient quality and integrity to perform genomic sequencing for 94% of the available samples (i.e., 612 out of 652 samples, Table 6). In contrast, there was a large failure rate of DNA extraction across all species for samples collected in the Moberly River in 2019 where out of 70 samples available, DNA was successfully extracted from only 32 (i.e., a 54% failure rate compared to a 0.3% failure rate when samples collected in the Moberly River in 2019 are excluded).

Preparation of the GBS library resulted in a pooled DNA library ready for sequencing with an average fragment size of 652 bp (coefficient of variation = 17.6%). Sequencing resulted in over 3.14 billion paired-end reads and demultiplexing resulted in the assignment of 2.63 billion paired-end reads to individual samples (i.e., 83.8% of reads were successfully assigned to a sample). Each sample had, on average, 8.12 million paired-end reads assigned (Table 7). This is a higher average number of reads assigned to a sample than for Bull Trout and Arctic Grayling, but lower than for Rainbow Trout.

Table 7. Comparison of GBS sequencing output (Number of reads per sample) for different species pooled libraries and different sequencing technologies.

	Bull Trout	Arctic Grayling	Rainbow Trout[a]	Non-game species[b]
Sequencing technology	HiSeq 4000	HiSeq 4000	NovaSeq6000 SP	NovaSeq6000 S4
Samples in pool	96	96	96	672

		Bull Trout	Arctic Grayling	Rainbow Trout[a]	Non-game species[b]
Reads per sample	Average	3,788,138	7,561,469	10,013,931	8,121,871
	Maximum	11,662,316	14,432,040	16,791,758	15,054,928
	Minimum	2,186,342	158,338	557,634	22,246

[a]For Rainbow Trout, two pools of 96 samples were sequenced. Values presented are for the two 96 sample pools.
[b]Non-game species refers to the GBS pool sequenced in 2022, with 299 Slimy Sculpin, 218 Redside Shiner, 95 Longnose Dace and additional samples from other species (see text for details).

Seven samples out of the 300 in the Slimy Sculpin and Coastrange Sculpin dataset had fewer than 1.3 million paired end reads, while the remaining samples had between 2.4 and 14.6 million reads. Regardless, we used all 300 samples for read mapping to the reference genome and polymorphism identification. We identified 2,479,808 putative genetic variants across all samples. At this stage, the same seven samples, including the Coastrange Sculpin, with fewer than 1.3 million reads, were eliminated from further analysis as they also had a) fewer than 1 million paired end reads mapping to the reference genome (versus 1.6 to 12.7 million for the remaining samples), b) low mean read depth (0.02 to 1.56X versus 3.10 to 16.90X for the remaining samples), and c) high levels of missing genotypes at putative genetic variants (99.4 to 78.6% versus 64.3 to 45.7% for the remaining samples). After eliminating those samples, we filtered our dataset to eliminate insertion/deletion polymorphisms (1.82 million SNPs remain), SNPs with observed heterozygosity over 0.6 (1.45 million SNPs remain), and loci with fewer than 70% of genotypes with a minimum genotype quality of 10 (i.e., genotype calling accuracy 90% or higher; 690,748 high quality SNPs remain). Finally, we eliminated SNPs with minor allele frequency (MAF) below 1% (162,869 SNPs remain) and retained only unlinked SNPs, i.e., SNPs that were not in strong LD (20,250 SNPs remain).

Results from a PCA (Figure 1; Appendix IV) on the SNP293 dataset with 20,250 SNPs revealed that the first axis of variation (explaining 18.2% of variation) separated three samples from Section 5 of the Peace River from all other samples, while the second axis (explaining 2.4% of variation) separated samples from the Moberly River from the remaining samples from the Peace River (Sections 3, 5, and 7). Further axes of variation explained less than 1% of variation each and did not separate samples according to sampling location or sampling year.

Results from the Admixture analysis were in agreement with the PCA. A model with three genetic groups (K=3) was the best fit to the data (had the lowest Cross-validation error; CVE) and identified the same three samples from Section 5 as one genetic group, samples from the Moberly River as a second genetic group and the remaining samples from the Peace River as a third genetic group (Figure 1; Appendix IV). These analyses suggest that the three samples from Section 5 are distinct from all others and given that the PC axis that separated them explains such a large proportion of variation (18.2%) we suspected that they may not be Slimy Sculpin, but possibly a different species (see below).

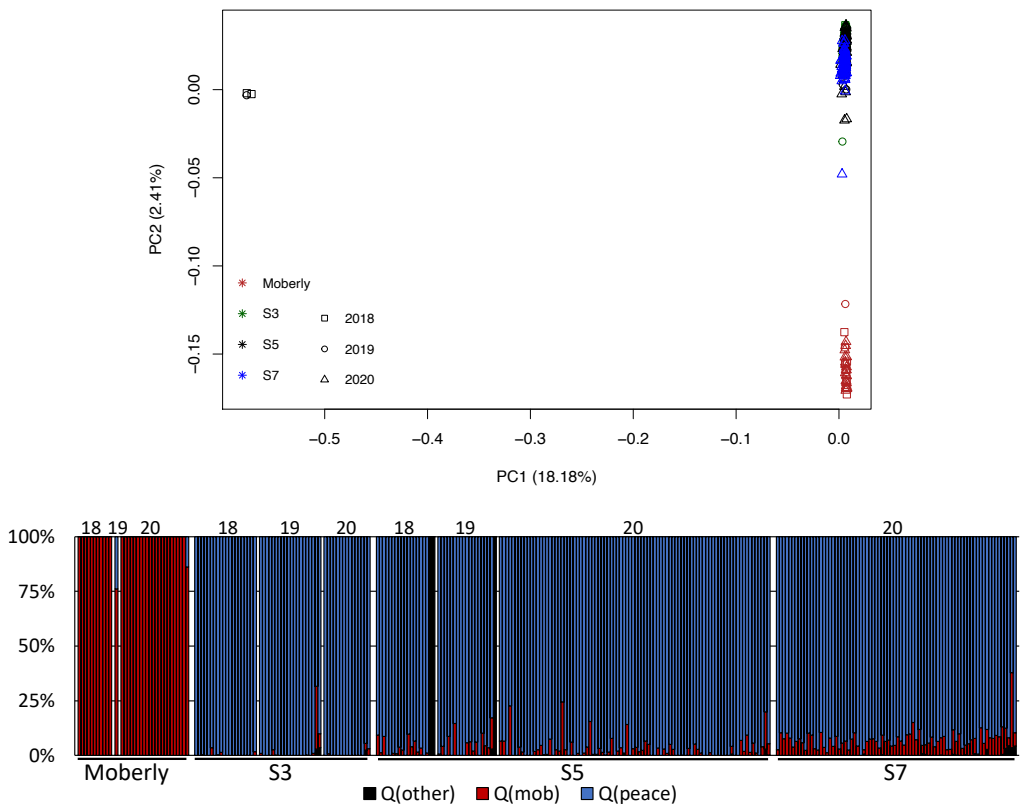

Figure 1. Population structure of Slimy Sculpin (SS) inferred with the SNP293 SS dataset (20,250 SNPs). Samples were collected in the Moberly River in 2018 (N=11), in 2019 (N=1) and in 2020 (N=22), in the Peace River Section 3 in 2018 (N=20), in 2019 (N=20) and in 2020 (N=15), Section 5 in 2018 (N=18), in 2019 (N=19) and in 2020 (N=89), and Section 7 in 2020 (N=22). The top panel shows the position of each sample along the first two axes of variation of a Principal Components Analysis. The sampling location is indicated by different colours (red for the Moberly River, green for Section 3, black for Section 5 and blue for Section 7) and the sampling year is indicated by the different symbols (square for 2018, circle for 2019 and triangle for 2020). The bottom panel shows the results of an Admixture analysis with three genetic groups. Each column represents the genotype of an individual fish, and the different colours represent the proportion of the genome of each fish that is assigned to each genetic (blue for the Peace River genetic group, red for the Moberly River genetic group and black for a group represented by three samples from the Peace River).

Results from the Admixture analysis were in agreement with the PCA. For the SNP290 dataset, a model with two genetic groups (K=2) was the best fit to the data (had the lowest CVE) and clearly separated samples from the Moberly River as one genetic group and the samples from the Peace River as a second genetic group (Figure 2; Appendix IV).

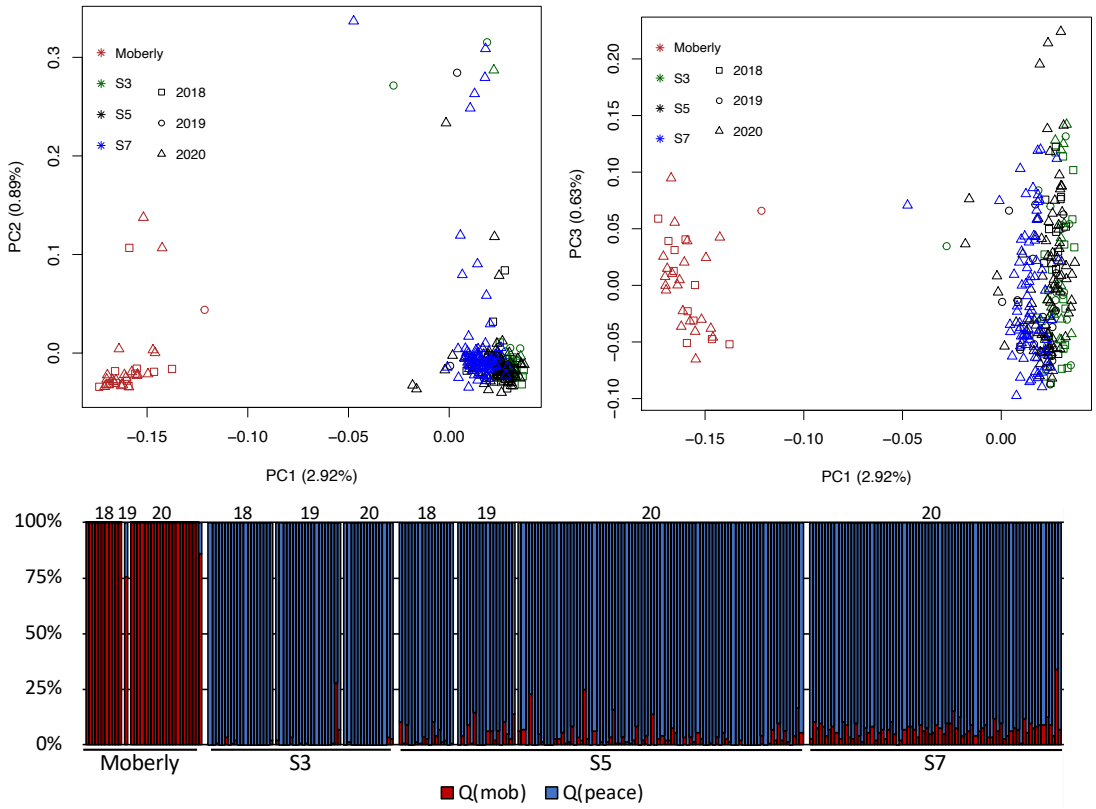

Figure 2. Population structure of Slimy Sculpin (SS) inferred with the SNP290 SS dataset (17,954 SNPs). Samples were collected in the Moberly River in 2018 (N=11), in 2019 (N=1) and in 2020 (N=22), in the Peace River Section 3 in 2018 (N=20), in 2019 (N=20) and in 2020 (N=15), Section 5 in 2018 (N=16), in 2019 (N=18) and in 2020 (N=89), and Section 7 in 2020 (N=22). The top left panel shows the position of each sample along the first two axes of variation of a Principal Components Analysis and the top right panel the position of along the first and third axes of variation. The sampling location is indicated by different colours (red for the Moberly River, green for Section 3, black for Section 5 and blue for Section 7) and the sampling year is indicated by the different symbols (square for 2018, circle for 2019 and triangle for 2020). The bottom panel shows the results of an Admixture analysis with two genetic groups. Each column represents the genotype of an individual fish, and the different colours represent the proportion of the genome of each fish that is assigned to each genetic (blue for the Peace River genetic group and red for the Moberly River genetic group).

No samples from either the Moberly River or the Peace River had a majority of their ancestry assigned to the opposite group (i.e., no sample caught in the Moberly had more than 50% ancestry in the Peace River group and vice versa) suggesting that migration between the two systems is limited. At the same time the data do suggest some admixture. Of note, 33.6% (86 out of 256) samples caught in the Peace River had more than 5% ancestry in the Moberly genetic group, with four of them having over 20%

ancestry in the Moberly group suggestive of them being early generation backcrosses to the Peace River group of fish. Evidence for admixture of Peace River ancestry into samples caught in the Moberly River is more limited. Only 2 (5.9%) of the 34 samples caught in the Moberly had ancestry in the Peace River genetic group (one had 24% and the other 14% ancestry in the Peace River group, again suggesting that they may be early generation backcrosses).

Finally, we quantified levels of genetic differentiation between sampling regions (Table 8) and between sampling years (Table 9) using the SNP290 dataset without eliminating linked loci.

Table 8. Weighted average Weir and Cockerham's F_{ST} between sampling regions of Slimy Sculpin in the LAA estimated for the SNP290 dataset with unlinked polymorphic loci with minor allele frequency above 1%.

	Moberly	Peace S3	Peace S5
Moberly			
Peace S3	0.108		
Peace S5	0.100	0.000	
Peace S7	0.093	0.001	0.001

In agreement with the PCA and Admixture analyses, there was no detectable genetic differentiation between different sampling sections of the Peace River (weighted average F_{ST} range 0.000-0.001; Table 8). By contrast, levels of genetic differentiation between the Moberly River and each of the three sampling regions of the Peace River were much higher: weighted average F_{ST} ranged from 0.093 to 0.108 (Table 8). Finally, there was no detectable genetic differentiation among years within each sampling region (Table 9).

Table 9. Weighted average Weir and Cockerham's F_{ST} between sampling regions and years of Slimy Sculpin in the LAA estimated for the SNP290 dataset with unlinked polymorphic loci with minor allele frequency above 1%. Highlighted cells indicate comparisons between years within each sampling region.

	Moberly 2018	Moberly 2020	Peace S3 2018	Peace S3 2019	Peace S3 2020	Peace S5 2018	Peace S5 2019	Peace S5 2020
Moberly 2018								
Moberly 2020	-0.001							
Peace S3 2018	0.110	0.114						
Peace S3 2019	0.106	0.109	0.001					
Peace S3 2020	0.112	0.115	0.000	0.000				
Peace S5 2018	0.104	0.107	0.001	0.000	-0.001			
Peace S5 2019	0.101	0.104	0.000	0.001	0.001	0.001		
Peace S5 2020	0.099	0.100	0.000	0.000	0.000	0.001	0.000	
Peace S7 2020	0.090	0.091	0.002	0.001	0.001	0.001	0.000	0.001

SPECIES DIAGNOSTICS

Materials and Methods

We took advantage of other GBS projects in the lab to generate SNP data for two ancillary Site C-related tasks aimed at determining the species identity of several fish whose identity was uncertain. For the first of two tasks, we wanted to determine if the three samples identified in the field as Slimy Sculpin (Table 10) but suspected to belong to a different species based on our analysis of population structure (see above), were actually Prickly Sculpin (*Cottus asper*) samples instead. Although the Spoonhead Sculpin (*C. ricei*) also occurs in the area it is much more rarely encountered and is morphologically quite distinctive and thus much less likely to be confused with Slimy Sculpin than is the Prickly Sculpin. To do so, we used GBS following the protocols detailed above to sequence three Prickly Sculpin samples collected in the LAA in a sequencing project independent from our Site C work. Demultiplexed reads from each sample were aligned to the genome reference sequence of *Cottus rhenanus* available online at Dryad (https://doi.org/10.5061/dryad.48g9f5r) following the protocols above. Variant identification was performed for each sample separately with GATK4 HaplotypeCaller (McKenna et al., 2010) and the results stored in individual Genomic Variant Call Format (GVCF) files. We then performed joint genotyping with the function GenotypeGVCFs after importing, into a Genomics Database with function GenomicsDBImport, the GVCF files for all 16 samples: a) the three new Prickly Sculpin samples, b) the three samples suspected of not being Slimy Sculpin, c) five Slimy Sculpin samples collected in the Moberly River, and d) five Slimy Sculpin samples from the Peace River. The resulting polymorphism file was then filtered to a) eliminate

insertion/deletion polymorphisms, b) eliminate SNPs with more than two variants, c) eliminate SNPs with observed heterozygosity above 0.6. We then prepared two separate datasets. One dataset was used to generate a phylogenetic network in SplitsTree4 (Huson and Bryant, 2006) after eliminating with VCFtools: a) sites with any level of missing data, b) sites with singletons (i.e., sites where the rare variant has a frequency below 1/32) and c) genotypes with genotype quality below 10. The second dataset was used to perform a PCA. This dataset is similar to the one used for the phylogenetic analyses except that we: a) allowed sites with up to 30% missing genotypes and b) we eliminated sites in close linkage.

Table 10. Samples used for the two Species Diagnostics tasks, the source of the sequencing data used (Sequencing), the sampling location of the samples (Location), the samples species as determined in the field (Field Species ID) and the samples species as determined by genetic analyses (Genetics Species ID).

UBC Code	Sequencing	Location	Field Species ID	Genetics Species ID
SS_S5_18_0746	Non-game species	Peace River - Section 5	Slimy Sculpin	Prickly Sculpin
SS_S5_19_0745	Non-game species	Peace River - Section 5	Slimy Sculpin	Prickly Sculpin
SS_S5_18_0512	Non-game species	Peace River - Section 5	Slimy Sculpin	Prickly Sculpin
PS-P3-18-141	Other Taylor Lab	Peace River - Section 3	Prickly Sculpin	
PS-P3-19-108	Other Taylor Lab	Peace River - Section 3	Prickly Sculpin	
PS-P5-19-540	Other Taylor Lab	Peace River - Section 5	Prickly Sculpin	
SS_MO_18_0045	Non-game species	Moberly River	Slimy Sculpin	
SS_MO_18_0051	Non-game species	Moberly River	Slimy Sculpin	
SS_MO_19_5637	Non-game species	Moberly River	Slimy Sculpin	
SS_MO_20_5783	Non-game species	Moberly River	Slimy Sculpin	
SS_MO_20_5786	Non-game species	Moberly River	Slimy Sculpin	
SS_S3_18_0402	Non-game species	Peace River - Section 3	Slimy Sculpin	
SS_S3_19_0141	Non-game species	Peace River - Section 3	Slimy Sculpin	
SS_S7_20_0560	Non-game species	Peace River - Section 7	Slimy Sculpin	

UBC Code	Sequencing	Location	Field Species ID	Genetics Species ID
SS_S7_20_0535	Non-game species	Peace River - Section 7	Slimy Sculpin	
SS_S5_20_0013	Non-game species	Peace River - Section 5	Slimy Sculpin	
EB-P5-21-069	Other Taylor Lab	Peace River - Section 5	Brook Trout or Brook Trout X Bull Trout Hybrid	Brook Trout
BR_UN_UN_102	Other Taylor Lab	Other	Brook Trout	
BT_BC_CB_197	Other Taylor Lab	Squamish River	Bull Trout	
BT_BC_CB_199	Other Taylor Lab	Squamish River	Bull Trout	
BT_BC_IB_200	Other Taylor Lab	Pine River - Wolverine River	Bull Trout	
BT_BC_IB_201	Other Taylor Lab	Halfway River - Fiddes Creek	Bull Trout	
LT_UN_UN_104	Other Taylor Lab	Other	Lake Trout	
AC_AC_AS_174	Other Taylor Lab	Canadian Arctic	Arctic Char	
DV_BC_AS_019	Other Taylor Lab	Stewart, BC	East Pacific Southern Dolly Varden	
DV_JP_AA_226	Other Taylor Lab	Hokkaido, Japan	West Pacific Southern Dolly Varden	
DV_RU_UN_066	Other Taylor Lab	Kamchatka, Russia	Northern Dolly Varden	

For the second task, one sample collected in Section 5 of the Peace River in 2021 (Table 10) could not be unambiguously identified in the field to the species level. The sampling crew suggested that it could be either a Brook Trout (*Salvelinus fontinalis*) or a Brook Trout by Bull Trout hybrid. To determine whether the sample is a Brook Trout or a Brook Trout by Bull Trout hybrid, we used GBS following the protocols detailed above to sequence (again, in a sequencing project independent from our Site C work) the sample in question and aligned the demultiplexed reads to the genome reference sequence of Dolly Varden (*Salvelinus malma*, assembly ASM291031v2; Christensen et al. 2018). Variant identification for this sample was performed as described above. We

then performed joint genotyping with the function GenotypeGVCFs, after importing into a Genomics Database with function GenomicsDBImport, the GVCF files for all 11 samples a) the potential Brook Trout or Brook Trout by Bull Trout hybrid from the LAA, b) one Brook Trout sample, c) two Bull Trout samples from the LAA, d) three Dolly Varden samples (one from British Columbia, one from Russia and one from Japan) e) one Arctic char (*Salvelinus alpinus*), and f) one Lake Trout (*Salvelinus namaycush;* all samples from Geraldes and Taylor, unpubl. data). Samples b) through f) served as reference samples to assist in identifying the unknown sample from the LAA. The resulting polymorphism file was then filtered to a) eliminate insertion/deletion polymorphisms, b) eliminate SNPs with more than two variants, c) eliminate SNPs with observed heterozygosity above 0.6, d) eliminate sites with any level of missing data, e) eliminate sites with singletons (i.e., sites where the rare variant has a frequency below 1/22), and f) eliminate genotypes with genotype quality below 10. This dataset was used to generate a phylogenetic network in SplitsTree4 (Huson and Bryant, 2006).

Results

We generated a phylogenetic network for the 16 sculpin samples with our filtered phylogenetic dataset (32,412 SNPs). The resulting network had a long internal branch separating two groups of samples which were all connected by very short branches (i.e., two genetically very divergent groups of closely related samples each; Figure 3). One group had all 10 samples from the Moberly River and Peace River genetic groups (identified in Figures 1 and 2) and the other group had the three Prickly Sculpin

samples, plus the three samples identified in the field as Slimy Sculpin which we suspected were from a different species (Figure 1).

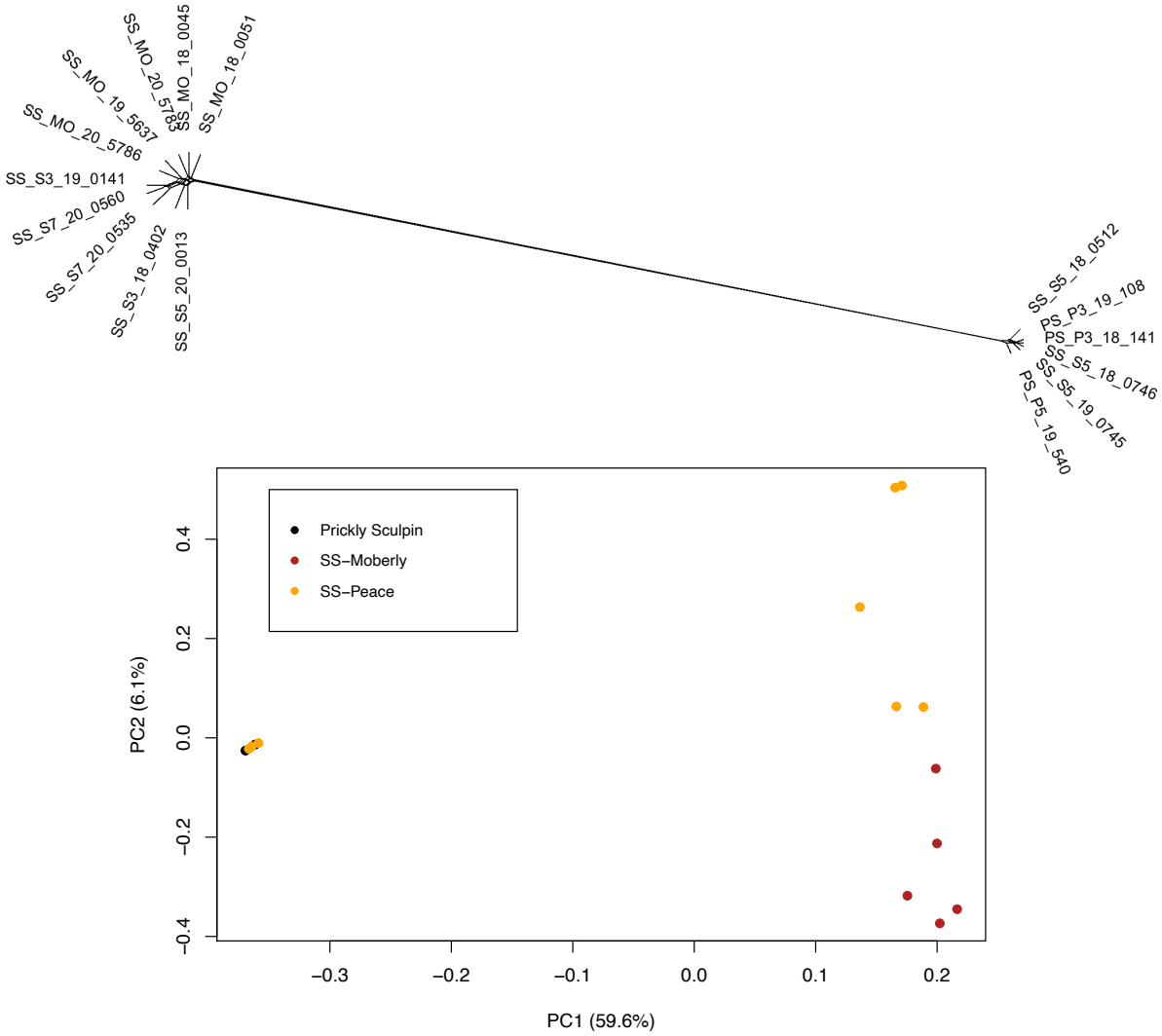

Figure 3. Phylogenetic Network (top panel; 32,412 SNPS) and Principal Components Analysis (bottom panel; 5,738) of 16 Sculpin samples. Both analyses included the three samples (SS_S5_18_0746, SS_S5_19_0745 and SS_S5_18_0512) identified in the field as Slimy Sculpin but suspected to belong to a different species based on population genetic analysis, as well as three Slimy Sculpin sample and 10 additional Slimy Sculpin samples from the Peace and Moberly rivers.

The PCA analysis (5,738 SNPs; Figure 3) revealed a similar pattern of two genetic groups separated along a first axis, which explained 59.6% of the variation, and again grouped to one side all 10 samples from the Moberly River and Peace River genetic groups (identified in Figures 1 and 2) and the other group had the three Prickly

Sculpin samples, plus the three samples identified in the field as Slimy Sculpin which we suspected were from a different species (Figure 1). The second axis, explaining only 6.1% of variation, separated the Moberly River and Peace River genetic groups. Taken together these analyses strongly suggest that the three samples identified as Slimy Sculpin, were in fact Prickly Sculpin.

To determine if one sample collected in Sections 5 of the Peace River in 2021 (Table 10) was either a Brook Trout or a Brook Trout by Bull Trout hybrid, we generated a phylogenetic network for 11 char (*Salvelinus* spp.) samples with our filtered phylogenetic dataset (51,407 SNPs). The resulting network clearly grouped together the reference Brook Trout sample and the sample suspected of being either Brook Trout or a hybrid with Bull Trout (Figure 4). These samples were separated by a long internal branch from all four Bull Trout samples in the analyses. Furthermore, because this is a phylogenetic network, if the sample were a Brook Trout and Bull Trout Hybrid, we would expect to see it emanating from a loop connecting the Brook Trout sample and the Bull Trout samples. This analysis strongly suggests that the sample collected in the LAA is a Brook Trout and not a Brook Trout by Bull Trout hybrid.

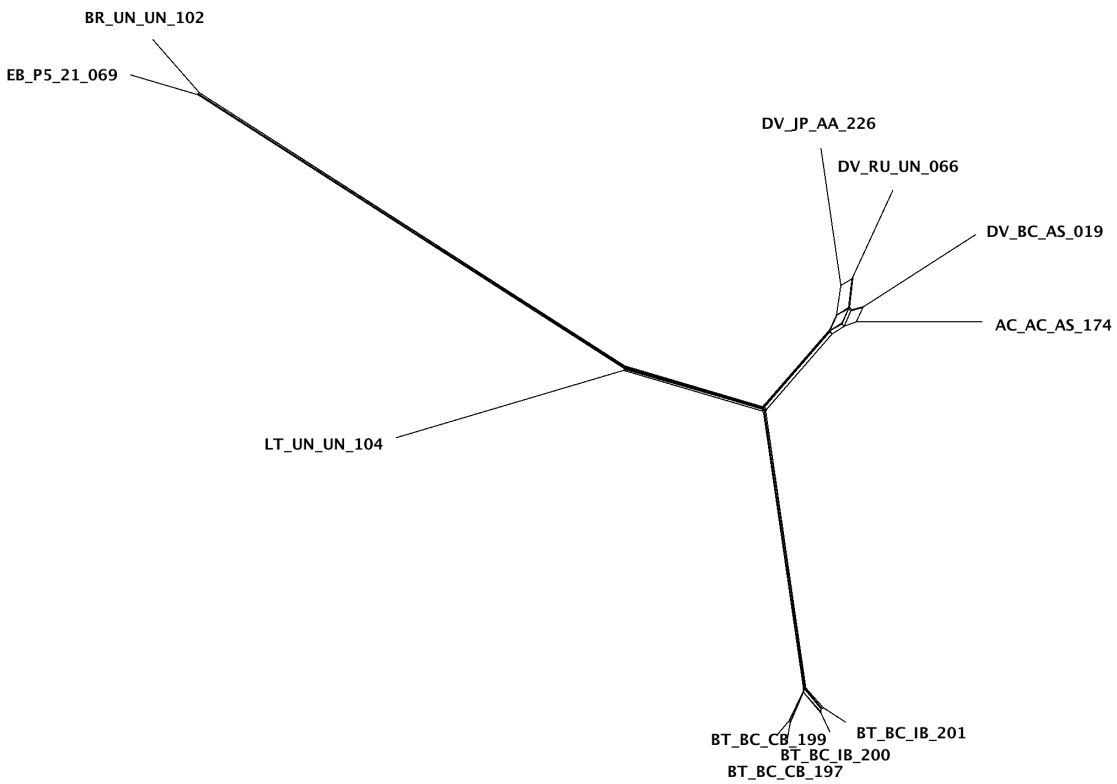

Figure 4. Phylogenetic Network (51,407 SNPs) of 11 char samples showed that sample EB_P5_21_069 is a Brook Trout sample. Analysis also included a Brook Trout sample (BR_UN_UN_102, a Lake Trout sample (LT_UN_UN_104), two interior (BT_BC_IB_201 and BT_BC_IB_200) and two coastal Bull Trout (BT_BC_CB_197 and BT_BC_CB_199) samples, one Arctic Char sample (AC_AC_AS_174) and three samples of Dolly Varden from different subspecies (DV_JP_AA_226, DV_RU_UN_066 and DV_BC_AS_019).

DISCUSSION

Our analyses of the samples collected in 2021 in the Peace River mainstem continue to find that the vast majority of Bull Trout, Arctic Grayling, and Rainbow Trout collected from throughout the various sampling sections of the Peace River mainstem originate from spawning tributaries upstream of the Project. The same pattern held true for the smaller number of fish of all three species samples within the TUF.

Furthermore, the assignment results for 2021 continue to highlight the importance of the Halfway River and the Moberly River as the most important tributaries for production of Bull Trout and Arctic Grayling, respectively, Consistent with results for previous years, assignment of Rainbow Trout to UP or DP produced the highest percentage of unassigned fish (e.g., four times that for Bull Trout), a result that Geraldes and Taylor (2022) suggested stemmed from the various impacts of stocking of hatchery strains within the LAA and adjacent areas (i.e., unknown status of individual populations as native or introduced, increasing straying of hatchery-produced fish and consequent genetic homogenization among populations).

For Bull Trout and Arctic Grayling, the predominance of fish assigned to UP occurs even when the majority of fish assayed in 2021 were collected from sections of the Peace River located wholly downstream of the Project (Sections 6, 7, and 9) and is similar to results from the 2016-2020 sample years. By contrast, there continues to be a higher proportion of Rainbow Trout assigned to DP, especially for fish collected from sections downstream of the Project. The higher proportion of fish assigned DP in 2021 may, in part, be driven by the fact that samples collected in sections of the Peace River located DP made up 12.7% of all samples collected in 2021 (19 out of 150 samples),

while in previous years only 4.0% of samples were collected DP (17 out of 427 samples). Overall, the use and production of Rainbow Trout seems to be more equitable between UP and DP portions of the LAA.

Our work expanded this year with the initial production of sequence data for three small-bodied non-game fish in the LAA: Slimy Sculpin, Longnose Dace, and Redside Shiner. These fishes belong to different families (Cottidae and Leuciscidae) than the game fishes previously studied by Geraldes and Taylor (2020, 2021, 2022), which suggests that the obvious differences in underlying biology (e.g., dispersal biology, reproductive biology, growth) compared to salmonid fishes may well have implications with respect to fish passage or their responses to habitat changes associated with the Project. In particular, the three non-game fishes are all characterized by much smaller maximum body sizes than the salmonids studied (generally less than 100 mm total length) and differences in body size and shape may impact swimming performance (e.g., Leavy and Bonner 2009). Some evidence exists that life-history may influence genetic diversity (e.g., Martinez et al. 2018) and that differences in dispersal biology can impact patterns of genetic structure across riverine fish species (e.g., Shelley et al. 2022). Furthermore, investigation across species representing diverse evolutionary histories and biology is a powerful approach to assess the generality of potential impacts of any development (e.g., Ruzich et al. 2019). Consequently, the inclusion of population genetic information of Redside Shiner, Longnose Dace, and Slimy Sculpin should provide a broader perspective to monitor impacts on fishes of the LAA.

The Slimy Sculpin is a common focus of investigation in environmental assessments owing to its broad geographic range in North America (from Alaska and

Yukon in the northwest to Virginia in the southeast, Scott and Crossman 1973; Gray et al. 2018). Recent molecular work, however, suggests that what is currently considered a single taxon, *Cottus cognatus*, may, in fact, consist of up to four distinct species (provisional species, PS, 29-30 and candidate species, CS, 31-32 of Young et al. 2022). Although samples from the Peace River were not examined by Young et al. (2022), it is likely that samples from the LAA belong to CS-31 and referred to as *C. cognatus* (sensu stricto) that were sampled from Alaska, Yukon and a portion of the Mackenzie River basin in southwest Northwest Territories (Young et al. 2022). Another candidate species, CS-32 was designated *C. philonips*, and was documented from the interior Columbia River basin and parts of Alaska. The two provisional species (i.e., those with no name proposed for them yet), 29 and 30, were found in eastern North America and western Washington, respectively, and remain unnamed (Young et al. 2022).

Our work found that there was no apparent genetic structure, when assessed by pairwise F_{ST}, among sample sites within the mainstem Peace River or between temporal samples within any single sample site (all $F_{ST} \sim 0.0$). By contrast, we found relatively striking and persistent genetic differences between the Moberly River sample site and all sites within the mainstem Peace River (max $F_{ST} = 0.114$, i.e., 115 times that among sites within the Peace River mainstem). In fact, the differences in F_{ST} between the Peace River mainstem and the Moberly River are comparable to those between Bull Trout from the Halfway and Pine rivers ($F_{ST} = 0.105$, Geraldes and Taylor 2020), and between Arctic Grayling from the Halfway, Moberly and Pine rivers from ($F_{ST} = 0.077 – 0.156$, Geraldes and Taylor 2021).

The apparent lack of differentiation of Slimy Sculpin within the Peace River mainstem is consistent with results from some other studies of Slimy Sculpin and in some other sculpin species, both in river and lake environments. Sample sites seem to be broadly interconnected genetically even with potential barriers to dispersal and in light of the generally benthic, lithophilic, and low dispersal biology of older juveniles and adult fish (e.g., Gray et al. 2018). At least some *Cottus* have, however, pelagic larval stages which may facilitate demographic exchange especially in the presence of strong water flows as in the Peace River mainstem (e.g., Dennenmoser et al. 2014). For instance, Euclide et al. (2018) documented a lack of genetic differentiation using microsatellite DNA between samples of Slimy Sculpin sampled from various parts of Lake Champlain, New York/Vermont. Sample sites were as much as 77 km apart from each other and at least partially isolated from each other by islands and causeways. There was also no difference reported between two sites in Lake Ontario separated from each other by 227 km (Euclide et al. 2018).

Notwithstanding these observations, there are sometimes pronounced genetic differences between population of freshwater sculpins in river environments (see Table 4 of Euclide et al. 2018, mean F_{ST} between populations ranged between 0.00 and 0.63 across nine species). Further, it is possible that subtle genetic differences exist within the mainstem Peace River in an isolation-by-distance (IBD) manner. Here, there are typically no striking differences among sites, but rather a more gradual, linear increase in genetic distance (e.g., F_{ST}) with increasing distance between sites. Such a pattern is often observed in marine fishes with a pelagic larval phase even though any individual pairwise F_{ST} measure is very low (i.e., < 0.01, e.g., Fitz et al. 2023). With additional

sampling across more sites of varying distances from one another it would be possible to test for IBD in the Slimy Sculpin. Perhaps 15 to 20 sites sampled across 100 km or more of river would provide suitable sampling for a robust test of this idea (e.g., Euclide et al. 2018 found no evidence of IBD across seven sites separated by as much as 77 km in Lake Champlain Slimy Sculpin).

These observations make the striking pattern of strong divergence between the Moberly and Peace rivers notable, especially given the presence of downstream flow from the former into the latter. One possibility is that environmental differences between the Moberly and the Peace rivers select for greater upstream rheotactic behaviour in Moberly River Slimy Sculpin such that they maintain occupancy in their natal systems which constrains gene flow between the rivers. Such differences in rheotaxis have been established in different populations across multiple species of salmonids (e.g., Raleigh 1971; Kaya 1989; Taylor 1988); the most striking example being biased upstream movement of Rainbow Trout fry in populations that spawn upstream of impassible waterfalls (Northcote 1981). It is also possible that Moberly River Slimy Sculpin do disperse widely downstream, but that such migrants are selected against in the Peace River mainstem.

Regardless of the reason for the strong differentiation between Slimy Sculpin from the Moberly and Peace rivers, it will be interesting to see if similar differences are observed in the Longnose Dace and Redside Shiner that have been sampled from the same areas. Over the next year we will be completing the analysis of those two species to assess this possibility. Finally, the strong differentiation between Slimy Sculpin from the Moberly and Peace rivers yields a clear signal that can be monitored over time as

the flow regime between the lower Moberly River and the Peace River mainstem changes with reservoir filling.

In conclusion, our work to date has resulted in genomic assays for efficient and accurate monitoring of population structure and for assignments of all three species to UP or DP and in some cases (Arctic Grayling) for assignment to tributary of origin. We have also resolved significant population structure in the Slimy Sculpin. In the coming months, assignments will continue for samples collected in 2022, and we will be continuing work we have started on: (i) developing more sensitive assays for Bull Trout and Rainbow Trout for the analysis of demographic characteristics (e.g., effective population size, genetic variation, parentage), (ii) population structure of two remaining non-game species (Longnose Dace and Redside Shiner), (iii) use existing data to develop more species diagnostic markers for Brook Trout and Lake Trout that are occasionally encountered in the LAA.

REFERENCES

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Research 19(9):1655–1664. http://doi.org/10.1101/gr.094052.109

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Mol Ecol, 22(11):3124-3140. https://doi.org/10.1111/mec.12354

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets, GigaScience, 4(1):s13742–015–0047–8. https://doi.org/10.1186/s13742-015-0047-8

Christensen KA, Rondeau EB, Minkley DR, Leong JS, Nugent CM, Danzmann RG, Ferguson MM, Stadnik A, Devlin RH, Muzzerall R, Edwards M, Davidson WS, Koop BF (2018) The Arctic charr (*Salvelinus alpinus*) genome and transcriptome assembly. PLoS One 13:e0204076. https://doi.org/10.1371/journal.pone.0204076

Retraction: Christensen KA, Rondeau EB, Minkley DR, Leong JS, Nugent CM, Danzmann RG, Ferguson MM, Stadnik A, Devlin RH, Muzzerall R, Edwards M, Davidson WS, Koop BF (2021) The Arctic charr (*Salvelinus alpinus*) genome and transcriptome assembly. PLoS One 16(2):e0247083. https://doi.org/10.1371/journal.pone.0247083

Huson DH, Bryant D (2006) Application of Phylogenetic Networks in Evolutionary Studies, Molecular Biology and Evolution, 23(2):254-267. https://doi.org/10.1093/molbev/msj030

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Durbin R (2011) The variant call format and VCFtools. Bioinformatics (Oxford, England), 27, 2156–2158. https://doi. org/10.1093/bioinformatics/btr330

Dennenmoser S, Sedlazeck FJ, Schatz MC, Altmüller J, Zytnicki M, Nolte AW (2019) Genome-wide patterns of transposon proliferation in an evolutionary young hybrid fish. Molecular Ecology, 28(6):1491–1505. https://doi.org/10.1111/mec.14969

Dennenmoser S, Rogers SM, Vamosi SM (2014) Genetic population structure in prickly sculpin (*Cottus asper*) reflects isolation-by-environment between two life-history ecotypes. Biological Journal of the Linnean Society 113:943-957. https://doi.org/10.1111/bij.12384

Euclide PT, Flores NM, Wargo MJ, Kilpatrick CW, Marsden JE (2018) Lack of genetic population structure of slimy sculpin in a large, fragmented lake. Ecology of Freshwater Fish 27:699-709. https://doi.org/10.1111/eff.12385

Fitz KS, Montes Jr HR, Thompson DM, Pinsky ML (2023) Isolation-by-distance and isolation-by-oceanography in Maroon Anemonefish (*Amphiprion biaculeatus*). Evolutionary Applications 16: 379-392. https://doi.org/10.1111/eva.13448

Geraldes A, Taylor EB (2020) Fish genetics study 2019 annual report. British Columbia Hydro Site C Clean Energy Report Series. https://sitecproject.com/sites/default/files/Fish-Genetics-Study-2019-Annual-Report.pdf. Accessed April 20, 2023.

Geraldes A, Taylor EB (2021) Fish genetics study 2020 annual report. British Columbia Hydro Site C Clean Energy Report Series. https://sitecproject.com/sites/default/files/Fish-Genetics-Study-2020-Annual-Report.pdf. Accessed April 20, 2023.

Geraldes A, Taylor EB (2022) Fish genetics study 2021 annual report. British Columbia Hydro Site C Clean Energy Report Series.

Gray MA, Curry RA, Arciszewski TJ, Munkittrick KR, Brasfield SM (2018) The biology and ecology of slimy sculpin: a recipe for effective environmental monitoring. Facets 3:103-127. https://doi.org/10.1139/facets-2017-0069

Irwin DE, Alcaide M, Delmore KE, Irwin JH, Owens GL (2016) Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. Mol Ecol 25:4488–4507. https://doi.org/10.1111/mec.13792

Kaya CM (1989). Rheotaxis of young Arctic grayling from populations that spawn in inlet or outlet streams of a lake. Transactions of the American Fisheries Society 118:474-481. https://doi.org/10.1577/1548-8659(1989)118<0474:ROYAGF>2.3.CO;2

Leavy TR, Bonner TH (2009) Relationships among swimming ability, current velocity association, and morphology for freshwater lotic fishes. North American Journal of Fisheries Management 29:72-83. https://doi.org/10.1577/M07-040.1

Martinez AS, Willoughby JR, Christie MR (2018) Genetic diversity in fishes is influenced by habitat type and life-history variation. Ecology and Evolution 8:12022-12031. https://doi.org/10.1002/ece3.4661

Northcote TG (1981) Juvenile current response, growth and maturity of above and below waterfall stocks of rainbow trout, *Salmo gairdneri*. Journal of Fish Biology 18:741-751. https://doi.org/10.1111/j.1095-8649.1981.tb03815.x

Owens GL, Baute GJ, Rieseberg LH (2016) Revisiting a classic case of introgression: hybridization and gene flow in Californian sunflowers. Mol Ecol 25:2630–2643. https://doi.org/10.1111/mec.13569

Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. Journal of Heredity, 95, 536-539. https://doi.org/10.1093/jhered/esh074

Raleigh RF (1971) Innate control of migrations of salmon and trout fry from natal gravels to rearing areas. Ecology 52:291-297. https://doi.org/10.2307/1934587

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proc Natl Acad Sci USA 94:9197–9201. https://doi.org/10.1073/pnas.94.17.9197

Ruzich J, Turnquist K, Nye N, Rowe D, Larson WA (2019) Isolation by a hydroelectric dam induces minimal impacts on genetic diversity and population structure in six fish species. Conservation Genetics 20:1421-1436. https://doi.org/10.1007/s10592-019-01220-1

Scott WB, Crossman EJ (1973) Freshwater fishes of Canada. Fisheries Research Board of Canada Bulletin 184. 966 p.

Shelley JJ, Holland OJ, Swearer SE, Dempster T, Le Feuvre MC, Sherman CD, Miller AD (2022) Landscape context and dispersal ability as determinants of population genetic structure in freshwater fishes. Freshwater Biology 67:338-352. https://doi.org/10.1111/fwb.13844

Taylor EB (1988) Adaptive variation in rheotactic and agonistic behavior in newly emerged fry of chinook salmon, *Oncorhynchus tshawytscha*, from ocean-and stream-type populations. Canadian Journal of Fisheries and Aquatic Sciences 45:237-243. https://doi.org/10.1139/f88-028

Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Young MK, Smith R, Pilgrim KL, Isaak DJ, McKelvey KS, Parkes S, Egge J, Schwartz MK (2022). A molecular taxonomy of *Cottus* in western North America. Western North American Naturalist 82:307-345.

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir B S (2012). A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. Bioinformatics. DOI: 10.1093/bioinformatics/bts606