Site C Clean Energy Project

Fisheries and Aquatic Habitat Monitoring and Follow-up Program

Fish Genetics Study
2023 Status Report for Bull Trout, Arctic Grayling, Rainbow Trout, and Longnose Dace

Construction Year 9 (2023)

Armando Geraldes, PhD
University of British Columbia

Eric B. Taylor, PhD
University of British Columbia

May 30, 2024

EXECUTIVE SUMMARY

BC Hydro is constructing the Site C Clean Energy Project (the Project) near the town of Fort St. John in northeastern British Columbia which will be the Peace River's third hydroelectric dam. 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 (TUF) have collected tissue 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*). The first phase of a Site C Fish Genetics Study was conducted between 2018 and 2021 by the laboratory of Eric Taylor at the University of British Columbia (UBC) where we: (a) determined levels and patterns of population structure in 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 was extended until the end of December 2025 with the following activities: 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, 2023 to December 31, 2023. Previous results and findings can be found in Geraldes and Taylor (2020, 2021, 2022, 2023).

For Activity 1, samples of Bull Trout, Arctic Grayling and Rainbow Trout for population assignment were collected in the Peace River in sampling year 2022 and 485 samples were received at UBC where they have been stored and catalogued.

For Bull Trout, 294 samples were collected in the Peace River in 2022, their DNA was extracted, and they were genotyped at six loci previously developed for population assignment to either of two genetic groups detected in the LAA: 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 294 Bull Trout samples collected in 2022 (including 15 sampled from the TUF), the vast majority of samples were assigned to UP (N=275, 93.5% of all samples) and a small number were assigned to DP (N=11, 3.7% of all samples). Of the 15 Bull Trout sampled at the TUF in 2022, none were assigned to DP and 13 (86.7%) were assigned to the UP group. Only eight (2.7%) of 294 fish could not be assigned to one of the two groups with more than 95% confidence.

For Arctic Grayling, 78 samples were collected in the Peace River in 2022, their DNA was extracted, and they were genotyped at 11 loci previously developed for population assignment (Geraldes and Taylor 2021). Geraldes and Taylor (2021) found that four distinct population groups of Arctic Grayling are found 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). All 78 fish samples, including the 46 sampled at the TUF, were assigned to the UP group. Of those, only five samples (6.4%) could not be subsequently assigned to a specific tributary. The remaining were assigned to the Moberly River group.

Finally, for Rainbow Trout, 107 samples were collected in the Peace River in 2022, their DNA was extracted, and they were genotyped at six loci previously developed for population assignment (Geraldes and Taylor 2022). 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), were identified. Of the 107 samples subject to assignment tests in 2022, 59 (55.1%) were assigned to the UP group, 24 (22.4%) to the DP group, and 24 (22.4%) could not be assigned with at least 95% confidence. No fish collected from the TUF were assigned to DP, four (66.7%) were assigned to UP, and two (33.3%) could not be assigned with more than 95% confidence to either group.

Activity 2 consists of the development and deployment of medium size SNP panels (200-300 loci) for Bull Trout and Rainbow Trout. The panels will allow for demographic inference with the goal of monitoring the effects of the Project on

populations of the two species. Work in 2023 focused on the development of the SNP panel for Bull Trout. Different iterations of the panel were tested with three rounds of genotyping in an initial set of samples (80 Bull Trout, four Arctic Char, three Dolly Varden, four Brook Trout and four Lake Trout). The optimized panel has a total of 190 loci that can be genotyped in thousands of samples at once. The panel includes one sex-linked locus for sex-identification, 17 species-specific loci that allow for the diagnosis of Bull Trout versus other *Salvelinus* char species (Arctic Char, Dolly Varden, Lake Trout and Brook Trout), 15 loci with large allelic differences between the UP and DP genetic clusters detected in the LAA and, 157 loci that show intermediate allele frequencies and little differentiation between the UP and DP genetic clusters that will be used for demographic inferences and monitoring the biological effectiveness of the TUF. Work on the development of a similar panel for Rainbow Trout will be performed in 2024.

An additional 1,023 samples of the three game species, collected in Peace River tributaries in the LAA in 2022, were received at UBC and catalogued. Among these samples, extraction and quality control of DNA were performed for 777 Bull Trout and 217 samples of Rainbow Trout; no additional samples of Arctic Grayling (29) were processed. The samples of Bull Trout and Rainbow Trout will be used for demographic inferences (Activity 2) in subsequent years.

For Activity 3, 888 samples collected in Peace River tributaries in the LAA in 2022 were received at UBC and catalogued. Samples were collected in the Peace River mainstem, the TUF and the Moberly River for a total of 758 samples of Redside Shiner,

72 samples of Slimy Sculpin and 54 samples of Longnose Dace. No DNA extractions for these three species were performed in 2023.

Geraldes and Taylor (2023) 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 of these three species, and examined population structure in Slimy Sculpin. Briefly, 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.

Over the past year we examined population structure of Longnose Dace based on 95 samples collected in the LAA, most were from the Moberly River (N=70), and the remaining were from the Peace River mainstem (N=23) and Maurice Creek (N=2). Additionally, two samples from southwestern BC were included. Surprisingly, a pattern of deep genetic divergence was detected where two samples from the Moberly River represented a lineage that was more distinct from the remaining samples from the LAA than were the two samples from southwestern BC. Four samples appeared to comprise genetic admixtures between the two lineages detected in the LAA. Work in 2024 will try to place the two lineages detected in the LAA into the wider pattern of genetic variability of the species in Canada. When those six samples are excluded, a subtle pattern of genetic differentiation was detected comprising two groups: one found mostly in the Peace River and the other mostly in the Moberly River.

Analysis of population structure in the LAA for Redside Shiner will be performed in 2024.

Finally, as in 2022 (Geraldes and Taylor 2023), we took advantage of other projects and our catalogue of genetic variants in *Salvelinus* spp. to identify to the species level three samples of char collected in the LAA. All three were determined to be Lake Trout (*Salvelinus namaycush*).

ACKNOWLEDGMENTS

A contract from BC Hydro supported this work. We thank Nich Burnett and Brent Mossop for continued guidance and helpful discussions. Dustin Ford of WSP, Mark LeRuez of Triton Environmental Consultants, and Palmer Environmental Consulting Group are thanked for assistance with sample collection, provision and orientation. Jessica Shen, Emma Laqua, and Brian Wu (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 Eric Taylor (Discovery and Equipment grant programs). The bioinformatics aspects of the project were supported by the UBC Department of Zoology's Computing Unit.

LIST OF ACRONYMS AND ABBREVIATIONS

ACDV	Arctic Char and Dolly Varden phylogenetic cluster
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 total variation (T)
GBS	Genotyping-by-sequencing
GT-seq	Genotyping-in-Thousands by sequencing
GVCF	Genomic Variant Call Format
HA	Halfway River population group of Rainbow Trout
IBD	Isolation-by-distance
K	Number of genetic groups in the Admixture analysis
LAA	Local Assessment Area
LD	Linkage Disequilibrium
MHC	Major Histocompatibility Complex
ML	Moberly River and Lynx Creek population group of Rainbow Trout
N_e	Effective Population Size
N_c	Census Population Size
PCA	Principal components analysis

PCR	Polymerase chain reaction
PR	Peace River
Q	Genomic proportion in Admixture analyses
QC	Quality control
SNP	Single nucleotide polymorphism
TUF	Temporary Upstream Fish Passage Facility
UBC	University of British Columbia
UP	Upstream of the Project

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Appendix 1. 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_BCHydroreport2024.xlsx"

Appendix 2. 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_BCHydroreport2024.xlsx"

Appendix 3. 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_BCHydroreport2024.xlsx"

Appendix 4. Longnose Dace library sample information, sequencing results, genomic coverage, frequency of missing genotypes, inclusion or not in the SNP96 and SNP83 datasets, PCA and Admixture results. File name: "Appendix4_LongnoseDace_GBS_BCHydroreport2024"

INTRODUCTION

BC Hydro is currently in the ninth 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. Between 2018 and 2021, BC Hydro and the laboratory of Eric Taylor at the University of British Columbia (UBC), Department of Zoology, engaged in a collaborative research 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 (see Geraldes and Taylor 2020, 2021, 2022).

In September 2021 a new four-and-one half year agreement between the lab of Eric Taylor and BC Hydro took effect. The agreement comprises three activities: (1) continue the population assignment work for Bull Trout, Arctic Grayling, and Rainbow Trout from 2021 sample years onwards, (2) 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) 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. This work illustrates BC Hydro's use of multiple lines of evidence to better understand the population structure, migration, and movement patterns of 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) 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) develop and deploy genotyping assays for genetic monitoring and demographic analysis of Bull Trout and Rainbow Trout in the LAA, and (3) determine levels and patterns of genetic structure of Slimy Sculpin, Redside Shiner and Longnose Dace prior to river diversion as a baseline for future monitoring.

Work Conducted Prior to 2023

Geraldes and Taylor (2020, 2021, 2022, 2023) reported on the results of the initial genetic work contributing to the FAHMFP, focusing on the use of 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, samples from the same three rivers plus the Beatton River were

examined. 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 upstream of the Project), 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) reported 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).

Overall, about 94% of the 1,206 Bull Trout were assigned to UP and about 3% to DP between 2016 and 2021; only about 3% of mainstem Peace River samples of Bull Trout could not be assigned to either the UP or DP spawning groups with more than 95% confidence.

Of the 266 Arctic Grayling sampled from the mainstem Peace River between 2016 and 2021, 93% were assigned to UP and about 6% to DP; about 1.5% of the Arctic Grayling samples could not be assigned to either the UP or DP spawning groups with more than 95% confidence. For Arctic Grayling, population assignment showed that about 87% of fish were assigned to the Moberly River (located UP), 5% 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). About 7% of Arctic Grayling could not be assigned to individual tributaries with over 95% confidence.

For Rainbow Trout, the GBS data of Geraldes and Taylor (2022) indicated a pattern of population structure where three groups were identified largely corresponding to i) samples collected in the Halfway River (HA, located UP), ii) samples collected in the Moberly River and Lynx Creek (ML, also located UP), and iii) samples collected from tributaries of the Pine River, Blind Creek and Burnt River (BB, located DP). The results suggested that there were much higher levels of interbreeding (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, located DP), Maurice Creek and Farrell Creek (Peace River tributaries, located UP) 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 assignment of fish to the UP and DP genetic groups. The majority of the 577 LAA samples of Rainbow Trout from 2018-2021 were assigned to UP (54% vs 29% DP), but there was a high percentage (about 17%) of Rainbow Trout samples that were not able to be assigned to UP or DP groups with 95% or higher confidence.

For Bull Trout and Arctic Grayling, 100% (N = 17 and N = 11 respectively) of fish collected from the TUF were assigned to UP in 2021. Only one Rainbow Trout was collected from the TUF in 2021 and it was assigned to DP. For all three species there was little variation in UP vs DP assignment among sample years.

In 2022, 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 of Slimy Sculpin, Longnose Dace and Redside Shiner, and examined population structure in Slimy Sculpin (Geraldes and Taylor 2023). Briefly, 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.

Work Conducted Over The Past Year (2023)

The current report summarizes the work performed in 2023 on 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 2022 and provides a summary for all sample years between 2016 and 2022.

For the demographic analyses within Activity 2, DNA extractions of Bull Trout and Rainbow Trout from all sampling sites in the LAA were completed. In addition, we leveraged our extensive catalog of genetic polymorphism of Bull Trout samples from the LAA (Geraldes and Taylor 2020), as well as data from closely related *Salvelinus* species (Geraldes et al. 2023; Taylor et al unpublished data and this report), to develop a genotyping panel to efficiently genotype all samples of Bull Trout collected in the LAA with the main goal of performing demographic analysis through kinship analysis and

linkage disequilibrium approaches, to assess the potential impacts of the Project and the TUF. Here, there are three main kinds of interrelated analyses and metrics that are usually considered important in demographic monitoring within a conservation genetics framework that could be assayed with such a genotyping panel: (i) genetic diversity, (ii) individual reproductive success, and (iii) effective population size (see Hohenlohe et al. 2021 for a recent review).

Genetic diversity encompasses measures such as heterozygosity (the probability that an individual carries two distinct alleles at a locus), the number of polymorphic loci, and the inbreeding coefficient (the probability that two individuals share alleles at a locus that are identical by descent – i.e., mating between close relatives). An understanding of genetic diversity and its monitoring through time are valuable because losses of genetic diversity may represent the loss of genetic variants that are associated with greater survival in current or future environments (e.g., at adaptive loci) or may signal declines in population size which may have deleterious demographic (e.g., Allee effects) and genetic (e.g., inbreeding depression) impacts. Individual reproductive success refers to the idea that certain breeding individuals may contribute more (or less) to a resultant juvenile cohort. Here, the assaying of many loci in both potential parents and resultant offspring combined with analyses such as cohort replacement rate or parentage-based tagging (Evans et al.206; Weigel et al. 2019; Horn et al. 2024) can facilitate the accurate assignment of juvenile fish to parents whose characteristics are known and may vary (size, migration timing, breeding location, handling history, etc.). In this way, the characteristics and numbers of successful parents can be monitored through time for signals of variation following environmental changes. Finally, and

related to the above measures, is effective population size (N_e). In a conservation context, N_e is a measure of the actual (or "effective") number of parents that contributed to the cohort of fish being assayed for a metric such as heterozygosity or inbreeding. It is often compared to the census population size (N_c; the number of reproductively mature *potential* parents) and is typically considerably smaller (e.g., 10%) than N_c owing to several factors in natural populations (e.g., variance in reproductive ability, changes in population size; see Reiman and Allendorf 2001 for a discussion involving Bull Trout). Effective population size and N_c both can be estimated using polymorphic genetic markers. In particular, N_e impacts the rate of the loss of genetic diversity per generation (increased loss with lower N_e) and thus the relative importance of factors impacting population persistence such as inbreeding depression and the ability to adapt to environmental change. Consequently, the marker panels developed for Bull Trout and Rainbow Trout should be an important tool in monitoring these species in the LAA through time.

In addition to loci designed for demographic analyses, a small number of additional loci was included in the genotyping panel with the further goals of quickly confirming if samples were Bull Trout (or a related *Salvelinus* species or hybrids between species) and assigning samples to the UP or DP Bull Trout population groups in the LAA. In addition, one sex-specific marker was included with the goal of determining the sex of each fish genotyped. The resulting panel has 190 loci and will be used in subsequent years in all samples of Bull Trout collected in the LAA. Work on a similar panel for Rainbow Trout will be performed in 2024.

For Activity 3, we report on the analysis of population structure of the samples of Longnose Dace with GBS data generated in 2022 (Geraldes and Taylor 2023).

Finally, we report on the generation and analysis of GBS data for molecular-based identification of three fish that were identified in the field as char but morphology-based identification to the species level was uncertain.

ACTIVITY 1: BULL TROUT

Materials and Methods

A total of 1071 Bull Trout genetic samples were collected from the LAA in 2022 (Table 1). Subsequent DNA extraction and quality control (QC) of all 1071 samples followed Geraldes and Taylor (2020). A total of 300 of these samples were used in population assignments (Activity 1); the 771 samples collected in the LAA outside the mainstem of the Peace River (Table 1) were also extracted and will be genotyped in the future with the SNP panel developed (see below) to monitor demographic parameters in Bull Trout populations of the LAA (Activity 2).

We used six TaqMan™ assays designed from the GBS data as described by Geraldes and Taylor (2020) to efficiently genotype six ancestry informative SNPs (i.e., loci showing large levels of genetic differentiation between UP and DP genetic groups) and assign 294 Peace River Bull Trout samples collected in 2022 in the Peace River mainstem, including from the TUF, two samples collected in the Moberly River, and four samples collected in Maurice Creek (full methods in Geraldes and Taylor 2020 and 2021). Briefly, using the analytical procedure of Rannala and Mountain (1997) as implemented in the program GeneClass2 (Piry et al., 2004), 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 genetic work for Study Year 2022 and across all Study Years (2016-2022). Indicated are numbers of samples received (UBC), with DNA extracted (DNA) and genotyped at ancestry informative SNPs (TaqMan).

			Study Years 2016-2022			Study Year 2022 Only		
Species	**Watershed**	**River/SectionID**	**UBC**	**DNA**	**TaqMan**	**UBC**	**DNA**	**TaqMan**
All	All	All	6989	6799	2752	1502	1473	485
Bull Trout	All	All	4585	4585	1588	1071	1071	300
Bull Trout	Peace River	TUF	32	32	32	15	15	15
Bull Trout	Peace River	Section 1	302	302	302	46	46	46
Bull Trout	Peace River	Section 3	459	459	459	95	95	95
Bull Trout	Peace River	Section 5	404	404	404	85	85	85
Bull Trout	Peace River	Section 6	163	163	163	26	26	26
Bull Trout	Peace River	Section 7	100	100	100	18	18	18
Bull Trout	Peace River	Section 9	40	40	40	9	9	9
Bull Trout	Halfway River	Chowade River	1276	1276	16	282	282	0
Bull Trout	Halfway River	Colt Creek	40	40	13	12	12	0
Bull Trout	Halfway River	Cypress Creek	1203	1203	13	354	354	0
Bull Trout	Halfway River	Fiddes Creek	489	489	12	122	122	0
Bull Trout	Halfway River	Halfway River	7	7	6	0	0	0
Bull Trout	Halfway River	Halfway River	1	1	0	1	1	0
Bull Trout	Halfway River	Turnoff Creek	40	40	4	0	0	0
Bull Trout	Moberly River	Moberly River	11	11	8	2	2	2
Bull Trout	Peace River	Dry Creek	10	10	10	0	0	0
Bull Trout	Peace River	Maurice	8	8	6	4	4	4
Arctic Grayling	All	All	708	518	389	107	78	78
Arctic Grayling	Peace River	TUF	57	57	57	46	46	46
Arctic Grayling	Peace River	Section 1	5	5	5	1	1	1
Arctic Grayling	Peace River	Section 3	103	103	103	5	5	5
Arctic Grayling	Peace River	Section 5	104	104	104	25	25	25
Arctic Grayling	Peace River	Section 6	42	42	42	0	0	0
Arctic Grayling	Peace River	Section 7	28	28	27	1	1	1
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
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	4	1	1	3	0	0
Arctic Grayling	Halfway River	Kobes Creek	3	0	0	0	0	0
Arctic Grayling	Moberly River	Moberly River	225	42	12	26	0	0

			Study Years 2016-2022			Study Year 2022 Only		
Species	**Watershed**	**River/SectionID**	**UBC**	**DNA**	**TaqMan**	**UBC**	**DNA**	**TaqMan**
Rainbow Trout	All	All	1696	1696	775	324	324	107
Rainbow Trout	Peace River	TUF	6	6	6	6	6	6
Rainbow Trout	Peace River	Section 1	293	293	293	48	48	48
Rainbow Trout	Peace River	Section 3	268	268	268	25	25	25
Rainbow Trout	Peace River	Section 5	80	80	80	28	28	28
Rainbow Trout	Peace River	Section 6	14	14	14	0	0	0
Rainbow Trout	Peace River	Section 7	21	21	21	0	0	0
Rainbow Trout	Peace River	Section 9	1	1	1	0	0	0
Rainbow Trout	Halfway River	Chowade River	21	21	14	0	0	0
Rainbow Trout	Halfway River	Colt Creek	233	233	12	81	81	0
Rainbow Trout	Halfway River	Cypress Creek	33	33	14	0	0	0
Rainbow Trout	Halfway River	Kobes Creek	316	316	11	73	73	0
Rainbow Trout	Halfway River	Fiddes Creek	1	1	0	0	0	0
Rainbow Trout	Peace River	Dry Creek	7	7	7	0	0	0
Rainbow Trout	Peace River	Farrell Creek	280	280	23	27	27	0
Rainbow Trout	Peace River	Maurice Creek	122	122	11	36	36	0

Results

In 2022, 279 Bull Trout were collected in six sections of the Peace River. An additional 15 samples were collected from the TUF, two in the Moberly River and, four in Maurice Creek (Tables 1 and 2; Appendix I). All 300 samples were successfully genotyped at six ancestry informative loci with TaqMan™ assays. As in previous years, most samples collected in the Peace River mainstem were assigned to the UP group (N=275, 93.5%), only 11 were assigned to the DP group (3.7% of all samples), and eight could not be assigned to either group (i.e., assignment probability to either was below 0.95; 2.7% of all samples). Overall, there was little variability in the proportion of fish assigned to UP and DP between 2022 and all previous years (2016 through 2021; Table 2). Two samples (13.3%) collected from the TUF could not be assigned to either group with assignment probability over 0.95 and the remaining 13 were assigned to UP (86.7%,

Table 2). All samples collected in the Moberly River (N=2) and Maurice Creek (N=4) were assigned to UP (Appendix I).

Table 2 Number of Bull Trout samples caught in the Peace River (PR) or the Temporary Upstream Fish Passage Facility (TUF) 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 SNPs.

Location	Year	Total	UP	DP	Unassigned[1]
All Samples	2022	294	275 (93.5%)	11 (3.7%)	8 (2.7%)
	2016-2021	1206	1130 (93.7%)	38 (3.2%)	38 (3.2%)
	All years	1500	1405 (93.7%)	49 (3.3%)	46 (3.1%)
PR Section 1	2022	46	42 (91.3%)	2 (4.3%)	2 (4.3%)
	2016-2021	256	248 (96.9%)	5 (2.0%)	3 (1.2%)
	All years	302	290 (96.0%)	7 (2.3%)	5 (1.7%)
PR Section 3	2022	95	90 (94.7%)	2 (2.1%)	3 (3.2%)
	2016-2021	364	338 (92.9%)	11 (3.0%)	15 (4.1%)
	All years	459	428 (93.2%)	13 (2.8%)	18 (3.9%)
PR Section 5	2022	85	82 (96.5%)	3 (3.5%)	0 (0.0%)
	2016-2021	319	290 (91.0%)	13 (4.0%)	16 (5.0%)
	All years	407	372 (92.1%)	16 (3.9%)	16 (3.9%)
PR Section 6	2022	26	22 (84.6%)	3 (11.5%)	1 (3.8%)
	2016-2021	137	126 (92.0%)	9 (6.6%)	2 (1.5%)
	All years	163	148 (90.8%)	12 (7.4%)	3 (1.8%)
PR Section 7	2022	18	18 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	82	80 (97.6%)	0 (0.0%)	2 (2.4%)
	All years	100	98 (98.0%)	0 (0.0%)	2 (2.0%)
PR Section 9	2022	9	8 (88.9%)	1 (11.1%)	0 (0.0%)
	2016-2021	31	31 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	40	39 (97.5%)	1 (2.5%)	0 (0.0%)
TUF	2022	15	13 (86.7%)	0 (0.0%)	2 (13.3%)
	2021	17	17 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	29	30 (93.1%)	0 (0.0%)	2 (6.9%)

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

ACTIVITY 1: ARCTIC GRAYLING

Materials and Methods

A total of 107 Arctic Grayling samples were collected in 2022 from the LAA (Table 1). Subsequent DNA extraction and QC of all 78 samples collected in the Peace River itself, including 46 from the TUF, followed Geraldes and Taylor (2020). Twenty-nine samples collected in Peace River tributaries (Table 1) were catalogued but were not extracted or analyzed.

We used the 11 TaqMan™ assays designed from the GBS work described by Geraldes and Taylor (2021) to genotype the 78 Arctic Grayling samples collected in 2022 from the Peace River and to assign them to UP or DP, as well as to each of the four spawning tributaries using the methods described above for Bull Trout (see also Geraldes and Taylor 2021).

Results

All samples were successfully genotyped at 11 ancestry informative loci with TaqMan™ assays and all 78 were assigned to the UP group (Tables 3 and 4; Appendix II). This was the first sampling year where no samples were assigned to the DP group (from 2016 to 2021 an average of 5.6%, i.e., 15 out of 266 fished analysed, were assigned to the DP group). We note that while between 2016-2021, 28% of samples (74 out of 266) were collected in Sections 6, 7 and 9 of the Peace River downstream of the Project, only 1.2% of samples (1 out of 78) were collected in those sections in 2022 (Table 3).

Table 3. Number of Arctic Grayling samples collected in the Peace River (PR), including the Temporary Upstream Fish Passage Facility (TUF), 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 SNPs.

Location	Year	Total	UP	DP	Unassigned[1]
All Samples	2022	78	78 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	266	247 (92.9%)	15 (5.6%)	4 (1.5%)
	All years	344	325 (94.5%)	15 (4.4%)	4 (1.2%)
PR Section 1	2022	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	4	3 (75.0%)	1 (25.0%)	0 (0.0%)
	All years	5	4 (80.0%)	1 (20.0%)	0 (0.0%)
PR Section 3	2022	5	5 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	98	98 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	103	103 (100.0%)	0 (0.0%)	0 (0.0%)
PR Section 5	2022	25	25 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	79	79 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	104	104 (100.0%)	0 (0.0%)	0 (0.0%)
PR Section 6	2022	0	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	42	32 (76.2%)	7 (16.7%)	3 (7.1%)
	All years	42	32 (76.2%)	7 (16.7%)	3 (7.1%)
PR Section 7	2022	1	1 (100.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	26	21 (80.8%)	4 (15.4%)	1 (3.8%)
	All years	27	22 (81.5%)	4 (14.8%)	1 (3.7%)
PR Section 9	2022	0	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	6	3 (50.0%)	3 (50.0%)	0 (0.0%)
	All years	6	3 (50.0%)	3 (50.0%)	0 (0.0%)
TUF	2022	46	46 (100.0%)	0 (0.0%)	0 (0.0%)
	2021	11	11 (100.0%)	0 (0.0%)	0 (0.0%)
	All years	57	57 (100.0%)	0 (0.0%)	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=5, 6.4% in 2022 and N=19, 7.1% in previous years, Table 4). All samples that could be assigned to one specific tributary (N=73) were assigned to the Moberly River population group (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 SNPs.

Location	Year	Total	HA	MO	PI	BE	Unassigned[1]
All Samples	2022	78	0 (0.0%)	73 (93.6%)	0 (0.0%)	0 (0.0%)	5 (6.4%)
	2016-2021	266	2 (0.8%)	231 (86.8%)	14 (5.3%)	0 (0.0%)	19 (7.1%)
	All years	344	2 (0.6%)	304 (88.4%)	14 (4.1%)	0 (0.0%)	24 (7.0%)
PR Section 1	2022	1	0 (0.0%)	1 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	4	0 (0.0%)	3 (75.0%)	1 (25.0%)	0 (0.0%)	0 (0.0%)
	All years	5	0 (0.0%)	4 (80.0%)	1 (20.0%)	0 (0.0%)	0 (0.0%)
PR Section 3	2022	5	0 (0.0%)	5 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	98	1 (1.0%)	91 (92.9%)	0 (0.0%)	0 (0.0%)	6 (6.1%)
	All years	103	1 (1.0%)	96 (93.2%)	0 (0.0%)	0 (0.0%)	6 (5.8%)
PR Section 5	2022	25	0 (0.0%)	22 (88.0%)	0 (0.0%)	0 (0.0%)	3 (12.0%)
	2016-2021	79	0 (0.0%)	76 (96.2%)	0 (0.0%)	0 (0.0%)	3 (3.8%)
	All years	104	0 (0.0%)	98 (94.2%)	0 (0.0%)	0 (0.0%)	6 (5.8%)
PR Section 6	2022	0	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	42	0 (0.0%)	31 (73.8%)	6 (14.3%)	0 (0.0%)	5 (11.9%)
	All years	42	0 (0.0%)	31 (73.8%)	6 (14.3%)	0 (0.0%)	5 (11.9%)
PR Section 7	2022	1	0 (0.0%)	1 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	26	0 (0.0%)	18 (69.2%)	4 (15.4%)	0 (0.0%)	4 (15.4%)
	All years	27	0 (0.0%)	19 (70.4%)	4 (14.8%)	0 (0.0%)	4 (14.8%)
PR Section 9	2022	0	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	2016-2021	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%)

Location	Year	Total	HA	MO	PI	BE	Unassigned[1]
TUF	2022	46	0 (0.0%)	44 (95.7%)	0 (0.0%)	0 (0.0%)	2 (4.3%)
	2021	11	1 (9.1%)	10 (90.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	All years	57	1 (1.8%)	54 (94.7%)	0 (0.0%)	0 (0.0%)	2 (3.5%)

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

ACTIVITY 1: RAINBOW TROUT

Materials and Methods

A total of 324 Rainbow Trout genetic samples were collected in 2022 from the LAA (Table 1). Subsequent DNA extraction and QC of all samples followed Geraldes and Taylor (2020). A total of 107 of these samples were used in population assignments (Activity 1); the 217 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 (Activity 2) that will be reported in 2024.

We used the six TaqMan™ assays described by Geraldes and Taylor (2022) to genotype the 107 Rainbow Trout genetic samples collected in 2022 from the Peace River, following the methods described above for Bull Trout and Arctic Grayling (see also Geraldes and Taylor 2022).

Results

In 2022, 107 Rainbow Trout were collected from the Peace River mainstem including six samples collected from the TUF (Tables 1 and 5; Appendix III). All 107 samples were successfully genotyped at six ancestry informative loci with TaqMan™ assays. More than half of all samples were assigned to the UP group (N=59, 55.1% of all samples; Table 5). Of the remaining 48 samples, half were assigned to the DP group (N=24, 22.4% of all samples) and the other half could not be assigned to either group (22.4%). These values are close to those of previous years and although a higher percentage of fish collected from Section 5 were assigned UP than in previous years, the increase was not significant (contingency, chi-square, P = 0.22). Of the six samples

collected from the TUF, four were assigned to the UP group and two could not be assigned to either group with more than 95% confidence (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 SNPs.

Location	Year	Total	UP	DP	Unassigned[1]
All Peace River	2022	107	59 (55.1%)	24 (22.4%)	24 (22.4%)
	2018-2021	577	314 (54.4%)	166 (28.8%)	97 (16.8%)
	All years	684	373 (54.5%)	190 (27.8%)	121 (17.7%)
PR Section 1	2022	48	23 (47.9%)	11 (22.9%)	14 (29.2%)
	2018-2021	245	137 (55.9%)	58 (23.7%)	50 (20.4%)
	All years	293	160 (54.6%)	69 (23.5%)	64 (21.8%)
PR Section 3	2022	25	14 (56.0%)	6 (24.0%)	5 (20.0%)
	2018-2021	243	151 (62.1%)	56 (23.0%)	36 (14.8%)
	All years	268	165 (61.6%)	62 (23.1%)	41 (15.3%)
PR Section 5	2022	28	18 (64.3%)	7 (25.0%)	3 (10.7%)
	2018-2021	52	24 (46.2%)	18 (34.6%)	10 (19.2%)
	All years	80	42 (52.5%)	25 (31.3%)	13 (16.3%)
PR Section 6	2022	0	0 (0%)	0 (0%)	0 (0%)
	2018-2021	14	0 (0.0%)	14 (100.0%)	0 (0.0%)
	All years	14	0 (0.0%)	14 (100.0%)	0 (0.0%)
PR Section 7	2022	0	0 (0%)	0 (0%)	0 (0%)
	2018-2021	21	2 (9.5%)	18 (85.7%)	1 (4.8%)
	All years	21	2 (9.5%)	18 (85.7%)	1 (4.8%)
PR Section 9	2022	0	0 (0%)	0 (0%)	0 (0%)
	2018-2021	1	0 (0.0%)	1 (100.0%)	0 (0.0%)
	All years	1	0 (0.0%)	1 (100.0%)	0 (0.0%)
TUF	2022	6	4 (66.7%)	0 (0.0%)	2 (33.3%)
	2021	1	0 (0.0%)	1 (100.0%)	0 (0.0%)
	All years	7	4 (57.1%)	1 (14.3%)	2 (28.6%)

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

ACTIVITY 2: DEMOGRAPHIC ANALYSES

Materials and Methods

For the demographic analyses within Activity 2, DNA extractions and quality control was performed for all Bull Trout (Table 1, N=4,585) and Rainbow Trout (Table 1, N=1,696) samples from all sampling sites in the LAA.

The aim of this activity is to generate genetic polymorphism data that can be leveraged to perform demographic monitoring of Bull Trout and Rainbow Trout populations. We decided to use a technique called Genotyping-in-Thousands by sequencing (GT-seq, Campbell et al. 2014) as an efficient approach to generate data through the monitoring period of the Project. This approach uses next-generation sequencing of multiplexed Polymerase chain reaction (PCR) products to generate genotypes from relatively small panels (50–500) of targeted SNPs for thousands of individuals in a single sequencing lane of a next generation sequencer (e.g. Illumina). The method relies on two PCR steps, the first adds Illumina sequencing primer sites to amplicons and the second adds unique barcode sequences and Illumina capture sites to amplicons. The resulting individually barcoded products for each sample can then be pooled together and sequenced. The method has been applied to several species including some salmonids, e.g. Pink Salmon, *Oncorhynchus gorbuscha* (May et al. 2023), Kokanee*, O. nerka* (Chang et al. 2021), Steelhead Trout, *O. mykiss* (Campbell et al. 2015), and Bull Trout (Bohling et al. 2021). These panels have been successfully used to address conservation and management questions in these species, including evaluation of trap-and-haul systems in dam management (Weigel et al. 2019) and estimation of relative fitness of wild and hatchery individuals (Shedd et al. 2021).

Here we aimed to develop a panel for each species that would serve four purposes: i) identify the sex of each sample, ii) confirm species identity, iii) assign ancestry to UP and DP genetic groups, and iv) perform demographic monitoring of each species in the LAA. Our work in 2022 focused on Bull Trout.

For the above purposes we relied on previously generated data. The sex-identification locus, purpose i, was provided by Nathan Campbell of GTseek LLC (the company we collaborated with for Activity 2) from previous work. For the remaining purposes we relied on GBS data generated in the Taylor Lab. Specifically, the data for species identification, purpose ii, is from Liu et al. (2023), Geraldes and Taylor (2023), and Taylor et al. (unpublished). For this set of markers to be included in the panel we selected as candidate loci reported in Liu et al. (2023) and added additional loci that exhibited fixed allelic differences between species, and were either in different chromosomes, or if in the same chromosome were at least 5 M bp away from each other. The data for purpose iii, UP and DP genetic group identification, and purpose iv, demographic monitoring, are from Geraldes and Taylor (2020). We filtered our catalog of potential genetic variants found in 92 samples from the LAA that included only SNPs with observed heterozygosity below 0.6 (485,022 potential SNPs) to retain only the 2,009 SNPs that a) had average coverage across all samples of at least 8 reads, b) had fewer than 30% missing genotypes in each of the two genetic groups (UP and DP), c) had minimum genotype quality of 20 or above (indicating accuracy of 99% or higher), d) had a minor allele frequency of 10% or higher in the 92 samples from the LAA, and e) were assigned to a non-sex-linked linkage group (i.e. they were not sex-linked and were not in unplaced scaffolds). From this group of SNPs, for the purpose of determining

ancestry in UP and DP, purpose iii, we included the loci we have been using in our TaqMan assays (Geraldes and Taylor 2020) and supplemented those with loci that had a Weir and Cockerham's F_{ST} (Weir and Cockerham, 1984) value between UP and DP of at least 0.5, resulting in a pool of 67 candidates. For the purpose of demographic inference, purpose iv, we aimed to select loci that were common in both groups (i.e., minor allele frequency of 10% or higher in each of UP and DP) and low genetic differentiation between groups (i.e. Weir and Cockerham's F_{ST} between UP and DP below 0.2), resulting in a pool of 1,060 candidate SNPs. To ensure that loci included in each category were largely independent, we required that candidates for the ancestry in UP and DP group were at least 1 M bp apart and that the candidates for demographic inference were at least 164 K bp apart by selecting only the highest quality variant when multiple candidates were within those distances.

This pool of candidate loci was sent to our collaborator Nathan Campbell of GTseek LLC for primer design with an optimized pipeline that screens primers for their ability to amplify unique on-target regions of the genome and that through in-silico testing do not interfere with other primers in the mix. The selected primer pairs then underwent three rounds of testing through PCR and sequencing with a set of 95 samples including: 80 Bull Trout from the LAA (32 belonging to the DP group and collected in different tributaries in the Pine River system and 48 belonging to the UP group and collected in different tributaries in the Halfway River system and the Moberly River), four Lake Trout, four Brook Trout, three Dolly Varden and four Arctic Char samples. Of those 95 samples, 40 were known males, 38 were known females and 17 were of unknown sex. The sequence data generated for each round of testing were

analyzed in collaboration with Nathan Campbell of GTseek LLC following a pipeline used in similar projects. This included genotype calling for each sample where it was assumed that a heterozygote genotype would have close to 50:50 reads supporting each of the two alleles and that homozygous genotypes would have most reads supporting only one allele. In each round of testing, loci were dropped if they a) did not amplify, b) interfered too much with the performance of other primers, c) amplified off-target, d) amplified multiple loci, e) amplified too much and took a disproportionate number of reads in the sequencing run, and f) failed to discriminate the two alleles at the locus. The resulting inferred genotypes from the third and final round of testing were used to test the quality of the data generated by the final panel in three ways. First, for the 80 Bull Trout samples, we compared the genotype at each locus generated by GT-seq and GBS for all loci for which we had genotypes generated by both platforms for at least 70 samples. Second, we compared the sex determined by the GT-seq sex locus and the sex determined either by field observations or PCR sex determination following the protocol of Yano et al. (2012). Third, we ordinated the genotype data from the 95 samples and 190 loci in “genotype space” using principal components analyses (PCA) with the R package SNPrelate (Zheng et al. 2012) to summarize the genetic data and verify that the samples clustered in genotype space according to species and UP/DP genetic groups.

Results

A total of 333 primer pairs were selected for three rounds of PCR and sequencing testing (Table 6). These targeted: one sex identification locus (purpose i), 21 species

specific loci (purpose ii), 26 UP/DP ancestry informative loci (purpose iii) and 285 loci for demographic inference (purpose iv). Target loci were distributed over all 36 autosomal linkage groups of the *Salvelinus* spp. genome reference sequence with an average of nine loci per linkage group (range one locus on the smallest linkage group, LG21, and 27 on the largest linkage group, LG20). Three rounds of testing with a set of 95 samples led to the elimination of 143 loci and the retention of 190 loci in the final panel (Table 6) distributed across all but one (LG21) autosomal linkage groups. The final set of 190 loci included: one sex identification locus (purpose i), 17 species specific loci (purpose ii), 15 UP/DP ancestry informative loci (purpose iii) and 157 loci for demographic inference (purpose iv).

Table 6. Genomic distribution of autosomal loci in the Bull Trout GT-seq panel. In each cell the number of loci included in the final panel is indicated (with the number in parentheses indicating the number of loci selected after primer in-silico testing). For each linkage group, the number of loci for each purpose (ii: Species Identity, iii:UP/DP Ancestry, iv: Demographic inference) as well as the overall (All) number are shown. The sex identification locus (purpose i) was provided by Nate Campbell and is not included here.

Chromosomal contig	Linkage Group	Length (bp)	Species Identity	UP/DP Ancestry	Demographic Inference	All
NC_036838.1	LG1	58,017,395	2 (2)	0 (1)	6 (10)	8 (13)
NC_036839.1	LG2	43,538,721	3 (4)	0 (0)	5 (7)	8 (11)
NC_036840.1	LG3	36,001,405	1 (2)	0 (0)	0 (1)	1 (3)
NC_036844.1	LG5	37,080,635	2 (2)	0 (0)	4 (7)	6 (9)
NC_036845.1	LG6.1	30,249,148	1 (2)	0 (0)	3 (7)	4 (9)
NC_036846.1	LG6.2	26,025,374	1 (1)	0 (0)	2 (3)	3 (4)
NC_036847.1	LG7	34,303,021	1 (1)	0 (0)	6 (10)	7 (11)
NC_036848.1	LG8	54,842,065	0 (0)	1 (2)	3 (8)	4 (10)
NC_036849.1	LG9	32,654,316	0 (0)	1 (1)	7 (10)	8 (11)
NC_036850.1	LG10	22,457,292	1 (1)	0 (0)	7 (10)	8 (11)
NC_036851.1	LG11	51,124,027	0 (0)	1 (3)	5 (7)	6 (10)
NC_036852.1	LG12	13,980,584	1 (1)	0 (0)	3 (3)	4 (4)
NC_036853.1	LG13	50,975,424	0 (0)	0 (0)	3 (11)	3 (11)
NC_036854.1	LG14	54,096,485	0 (0)	2 (3)	8 (17)	10 (20)
NC_036855.1	LG15	67,329,100	0 (0)	0 (0)	4 (18)	4 (18)
NC_036856.1	LG16	42,871,064	1 (1)	1 (2)	6 (11)	8 (14)
NC_036857.1	LG17	41,841,263	0 (0)	0 (0)	6 (11)	6 (11)

Chromosomal contig	Linkage Group	Length (bp)	Species Identity	UP/DP Ancestry	Demographic Inference	All
NC_036858.1	LG18	72,741,121	0 (0)	1 (1)	5 (9)	6 (10)
NC_036859.1	LG19	38,228,754	0 (0)	0 (0)	3 (5)	3 (5)
NC_036860.1	LG20	79,996,362	0 (0)	1 (1)	14 (26)	15 (27)
NC_036861.1	LG21	6,905,391	0 (0)	0 (0)	0 (1)	0 (1)
NC_036862.1	LG22	37,604,395	0 (0)	1 (2)	1 (3)	2 (5)
NC_036863.1	LG23	49,632,736	1 (1)	0 (0)	11 (14)	12 (15)
NC_036864.1	LG24	11,432,800	0 (0)	0 (0)	1 (3)	1 (3)
NC_036865.1	LG25	26,198,113	0 (0)	1 (1)	4 (4)	5 (5)
NC_036866.1	LG26	49,931,436	0 (0)	0 (1)	4 (7)	4 (8)
NC_036867.1	LG27	38,733,064	0 (0)	0 (0)	6 (7)	6 (7)
NC_036868.1	LG28	32,734,159	0 (0)	3 (3)	3 (5)	6 (8)
NC_036869.1	LG30	26,193,892	0 (0)	0 (0)	2 (5)	2 (5)
NC_036870.1	LG31	32,006,513	0 (1)	1 (1)	5 (8)	6 (10)
NC_036871.1	LG32	38,480,802	1 (1)	1 (1)	2 (6)	4 (8)
NC_036872.1	LG33	38,084,510	0 (0)	0 (0)	7 (11)	7 (11)
NC_036873.1	LG34	8,958,605	0 (0)	0 (1)	1 (3)	1 (4)
NC_036874.1	LG35	21,595,701	0 (0)	0 (1)	2 (3)	2 (4)
NC_036875.1	LG36	41,232,801	0 (0)	0 (1)	5 (10)	5 (11)
NC_036876.1	LG37	19,546,989	1 (1)	0 (0)	3 (4)	4 (5)
All		**1,367,625,463**	**17 (21)**	**15 (26)**	**157 (285)**	**189 (332)**

For 93 out of the 95 samples included in the test plate, we were able to genotype 80% or more of the 190 loci in the GT-seq panel (for the remaining two Bull Trout samples, genotyping rates were only 21 and 10% suggesting that DNA quality or quantity may have caused the low genotyping rate). The genotyping rate was above 95% for 81 of the remaining 93 samples.

At 151 of the 190 loci included in the Bull Trout GT-seq panel, the genotypes inferred from GBS and GT-seq could be compared. Across all 151 loci, the concordance between them was on average 97.4%, and was above 98% for more than half the loci. The concordance was below 90% at only two of the 151 loci compared. Future adjustments to the genotype calling pipeline will likely improve genotype call agreement.

The GT-seq sex locus produced a genotype for 93 of the 95 samples in the test plate resulting in the identification of 43 female and 50 male samples. Of the 38 samples previously identified as female, 36 had a female genotype with the GT-seq locus and two did not produce a genotype. Of the 40 samples previously identified as male, 39 had a male genotype with the GT-seq locus and one had a female genotype. Overall, the concordance between the two assessments of sex was 98.7%.

Bull Trout samples are clearly separated from samples from all other species along the first PC (explaining 24.9% of the variation in the data) of the PCA of the GT-seq genotype data (Figure 1). The second PC (explaining 7.4% of the variation in the data) separates samples from the UP and DP genetic groups found in the LAA. The two DP samples closest to the UP samples along PC2 are not necessarily admixed between the two groups. Their position close to the origin of both PCs simply indicates that there is not enough genotype data for them in the dataset (these are the two samples with data at only 10 and 25% of 190 loci in the GT-seq panel). The separation along PC1 is most likely derived from the genotypes at the 17 species specific loci, and PC2 from the 15 UP/DP ancestry loci. The genotypes at the remaining 157 autosomal loci will be used for demographic inference in this system.

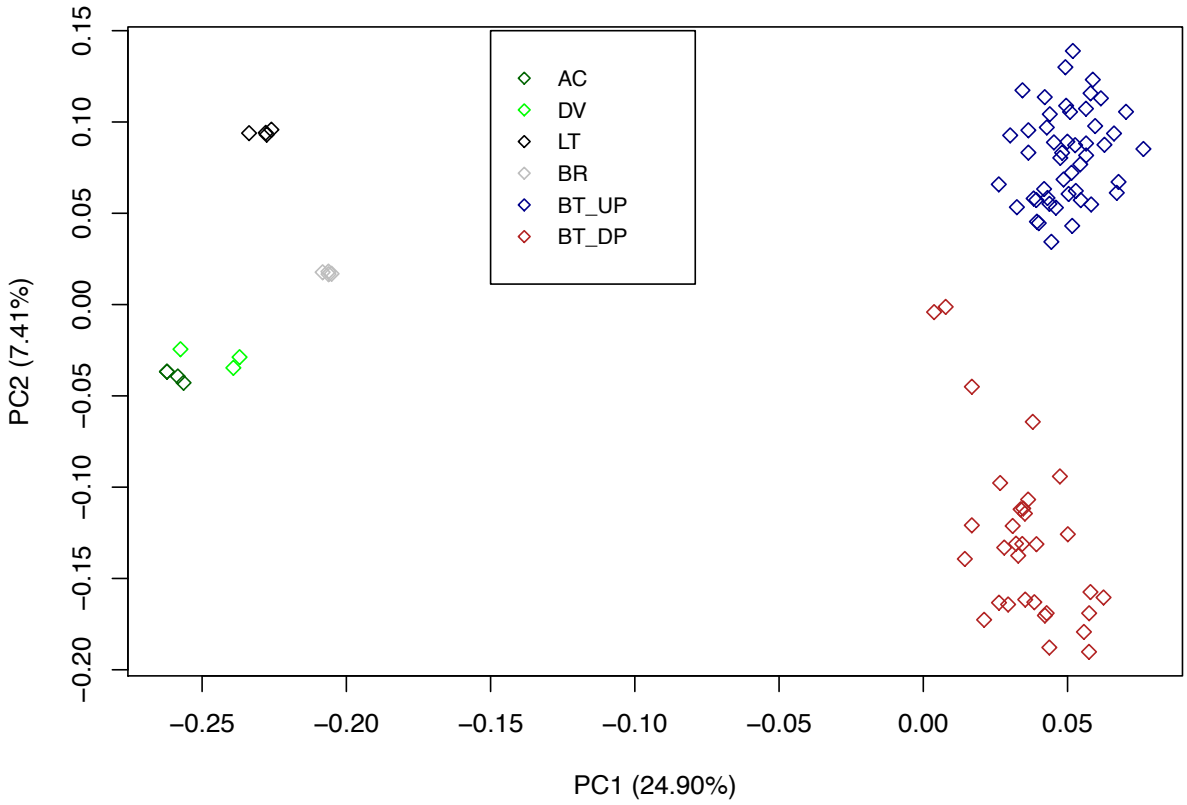

Figure 1. The first two Principal Components of a PCA with the Bull Trout GT-seq panel generated genotype data (190 loci). Samples are plotted as diamonds the colour of which indicate their species or genetic groups as indicated by the inset: Arctic char (N=4, dark green), Dolly Varden (N=3, light green), Lake Trout (N=4, black), Brook Trout (N=4, grey), Bull Trout from the Halfway and Moberly rivers (UP genetic group, N=48, blue), and Bull Trout from the Pine River (N=32, red). The amount of variation explained by each component is shown along each axis.

ACTIVITY 3: LONGNOSE DACE

Materials and Methods

The data for analysis of population structure of Longnose Dace was generated in 2022 in a large genotyping-by-sequencing (GBS) library (Geraldes and Taylor 2023) that included samples of the three non-game fish species (Slimy Sculpin, Redside Shiner and Longnose Dace).

Samples

DNA extraction and QC of all 103 samples of Longnose Dace sampled up to 2020 and received at UBC for genetic analysis followed Geraldes and Taylor (2020). Eight samples failed QC and were not selected for DNA sequencing (Table 7). We also included one Longnose Dace, and one sample of Nooksack Dace, a divergent form of Longnose Dace, from southwestern BC (both from the Chilliwack River system), for a total of 97 samples, 95 of which are from the LAA (Table 7).

Table 7. Number of samples of Longnose Dace 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).

River/SectionID	Year	UBC[1]	GBS[2]	SNP96[3]	SNP83[4]
All	All	103	95	94	83
Peace River-Section 3	2019	3	3	3	2
Peace River-Section 5	2019	5	5	5	3
Peace River-Section 5	2020	7	7	7	6
Peace River-Section 7	2020	8	8	8	7
Moberly River	2018	20	20	20	20
Moberly River	2019	24	16	16	13
Moberly River	2020	34	34	33	30
Peace River-Maurice Creek	2006	2	2	2	2

[1]Number of samples from the LAA received at UBC
[2]Number of samples from the LAA used for genotyping-by-sequencing (GBS)
[3]Number of samples from the LAA used for population genetic analysis
[4]Number of samples from the LAA used for population genetic analysis after eliminating 11 samples that were identified as being from a divergent lineage or admixed with that lineage (see text for details)

Sequencing, read mapping and variant identification

We used reduced representation genomic DNA sequencing with genotyping-by-sequencing (GBS) for sequence data generation and genetic variant discovery (single nucleotide polymorphisms, SNPs). Detailed descriptions of library preparation and sequencing were reported by Taylor and Geraldes (2023). The DNA library was sequenced in an Illumina NovaSeq 6000 S4 with 150 bp paired end reads at the McGill University and Génome Québec Innovation Centre in 2022.

For Longnose Dace 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 study (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), the Rainbow Trout (Geraldes and Taylor 2022) and Slimy Sculpin datasets (Taylor and Geraldes 2023). For the pooled DNA libraries sequenced in 2022, we used dual barcoding, i.e. each sample is barcoded with a combination of a well and a plate barcode. Reads were demultiplexed and assigned to individual samples with the function “process_radtags” from the STACKS v2.5 pipeline (Catchen et al. 2013) by analysing the two barcodes present, one in each of the two paired reads.

After demultiplexing, the sequence reads for Longnose Dace were aligned to the genome reference sequence of the Speckled Dace (*Rhinichthys ocsulus*), a closely related species (estimated divergence time is about 6 MYA). This genome was sequenced as part of the California Conservation Genomics Project (ccgproject.org) and was generated with PacBio long read data and Dovetail OmniC libraries, and Illumina sequencing at a University of California Sequencing Centre and is available

online at (https://ncbi.nlm.nih.gov/bioproject/PRJNA765855). This pseudo haplotype scaffold-level genome assembly includes 490 scaffolds and no assembled chromosomes.

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), Rainbow Trout (Geraldes and Taylor 2022) and Slimy Sculpin (Geraldes and Taylor 2022) datasets. One sample from the Moberly River (dataset SNP96; Table 7 and Appendix 4) was eliminated from the analysis prior to polymorphism identification as it generated less than 50,000 reads (average number of reads for the remaining samples was 8.6 million reads, range 1.5 to 14 M reads).

Analyses of Population Structure in Longnose Dace

After polymorphism identification in the remaining 96 samples (SNP96 dataset; Table 7) 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, and iii) SNPs with genotype quality below 10 (these have a higher than 10% chance of being incorrect genotypes).

At this stage, we further filtered this SNP96 dataset in two ways. To generate a phylogenetic network with SplitsTree4 (Huson and Bryant, 2006), we also eliminated from the SNP96 dataset any sites with missing genotypes and sites where the rare

variant was only observed once. For population genetic analyses we instead filtered the SNP96 dataset to eliminate loci with missing genotypes in more than 30% of samples, and low frequency SNPs (SNPs present at a frequency below 1%). We then used Plinkv1.9 (Chang et al. 2019) to remove SNPs that were in close linkage disequilibrium (LD) 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.

We used two complementary and independent approaches to infer patterns of population structure in Longnose Dace. 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 two

last runs (one with K of 2 and one with K of 3) using 1,000 bootstraps to estimate the standard error of the inferred admixture proportions for each K.

Analysis of population structure suggested that two samples from the Moberly River belong to a highly divergent lineage of Longnose Dace. An additional nine samples have evidence of some admixture between the most common lineage in the LAA and that divergent lineage. To further dissect patterns of population structure within this most common Longnose Dace lineage in the LAA with PCA and Admixture analyses, we generated a second SNP dataset (SNP83; Table 7) that includes only the 83 samples from the LAA that show little or no evidence of admixture with the divergent lineage of Longnose Dace (Q<0.02 in the Admixture analysis with K=3 in the SNP96 dataset analysis; see Results section below). We then repeated the filtering described above for analyses of population structure for this SNP83 dataset.

Finally, 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 for those sampling regions where at least five samples remained in this SNP83 dataset. This analysis was performed for all SNPs that remained after filtering for population genetics analysis but prior to LD pruning.

Results

We identified 2,050,342 putative genetic variants across all Longnose Dace samples in the SNP96 dataset. We filtered that dataset to eliminate insertion/deletion polymorphisms (1.57 million SNPs remain) and SNPs with observed heterozygosity over 0.6. The resulting SNP96 dataset retained 1.02 million (M) SNPs.

We then kept 7,137 sites for the phylogenetic network analysis in SplitsTree4 (Huson and Bryant, 2006) after eliminating from the SNP96 dataset genotypes with Genotype Quality below 10, sites with missing genotypes and genotypes where the allele count of the rare allele was only one. The resulting phylogenetic network has a long internal branch separating two groups of samples (Figure 2): one group of very closely related fish contains most samples from the LAA and two samples of Longnose Dace from southwestern BC, the other group has two samples from the Moberly River (indicated with black numbers in Figure 2). Four samples from the LAA (indicated with orange numbers in Figure 2) appear as reticulations (net-like connections) between the two groups along the long internal branch. These results suggest that there are two genetically divergent groups of Longnose Dace in the LAA, one that is closely related to samples from southwestern BC and one that is highly divergent from these samples and others from the LAA. It furthermore suggests that some samples are hybrids between these two divergent lineages.

For population genetic analyses, the SNP96 dataset was filtered down to 26,074 unlinked SNPs with less than 30% missing genotypes, and minor allele frequency of at least 1% among the 96 samples. Results from a PCA (Figure 3; Appendix IV) on the SNP96 dataset revealed a similar pattern to the phylogenetic network. The first axis of variation (PC1, explaining 7.7% of variation) separated two samples collected in the Moberly River from most other samples from the LAA, while the second axis (PC2, explaining 5.9% of variation) separated the Nooksack Dace and the Longnose Dace from southwestern BC from samples from the LAA. The same four samples that appeared as reticulations in the phylogenetic network (indicated with orange numbers in

Figure 2) appear intermediate along PC1 suggesting that they are admixed between the two lineages of Longnose Dace detected here.

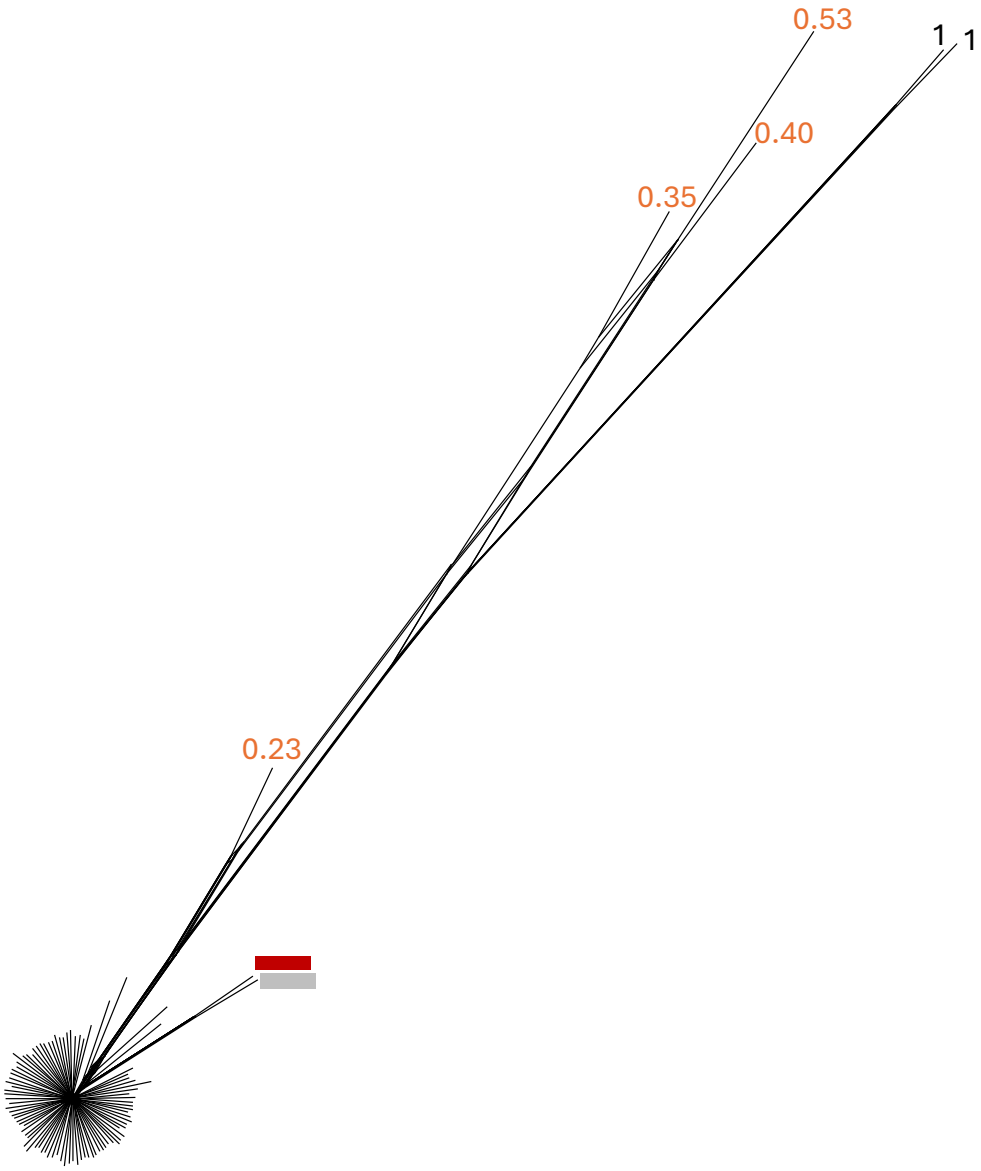

Figure 2. Phylogenetic Network using the SNP96 dataset for phylogenetic analysis of Longnose Dace (7,137 SNPS). All samples are from the LAA except for two samples which are from southwestern BC (a Longnose Dace indicated by a red box and one Nooksack Dace indicated by a grey box). Numbers indicate ancestry proportions (Q value) from the Admixture analysis with K=2 (see below) for six samples, numbers in black for two samples that represent a divergent lineage and numbers in orange for four samples that appear as reticulations between the two lineages detected.

A third principal component axis of variation (PC3) explained 1.6% of variation and separated the Nooksack Dace from the southwestern BC Longnose Dace (Appendix IV), a fourth axis (PC4) explained 1.4% of the variation and separated a single sample from the Moberly River (Appendix IV). Finally, a fifth principal (PC5) component explaining 1.4% of variation imperfectly separated samples from the Moberly River from samples from the Peace River mainstem (Appendix IV). In no case did we observe a separation of samples according to sampling year.

Results from the Admixture analyses agreed with the PCA (Figure 3; Appendix IV). A model with two genetic groups (K=2) was the best fit to the data (had the lowest Cross Validation Error; CVE) and identified the same two samples from the Moberly River as one genetic group, and most of the remaining samples from the LAA as a different genetic group. Nine samples had more than 2% of their inferred ancestry from the divergent genetic group of Longnose Dace. The four samples that appear as reticulations in the phylogenetic network (Figure 2) were the only samples to have more than 22% of their genome as admixed, the remining five had admixture estimates between 9 and 13% (Figure 3 and Appendix IV). Admixed samples were present in all sampling locations and years except for Maurice Creek and Moberly River in 2018. The two dace samples from southwestern BC appear as admixed between the two genetic groups. Those samples become their own genetic group when K=3 is used (Figure 2C). In that analysis, 12 samples from the LAA appear as admixed with this third genetic group (ancestry proportion range in that third group between 4 and 17%). Individual samples ancestry proportions from the divergent genetic lineage identified in PC1 and the phylogenetic network is largely similar for analyses with K=2 and K=3.

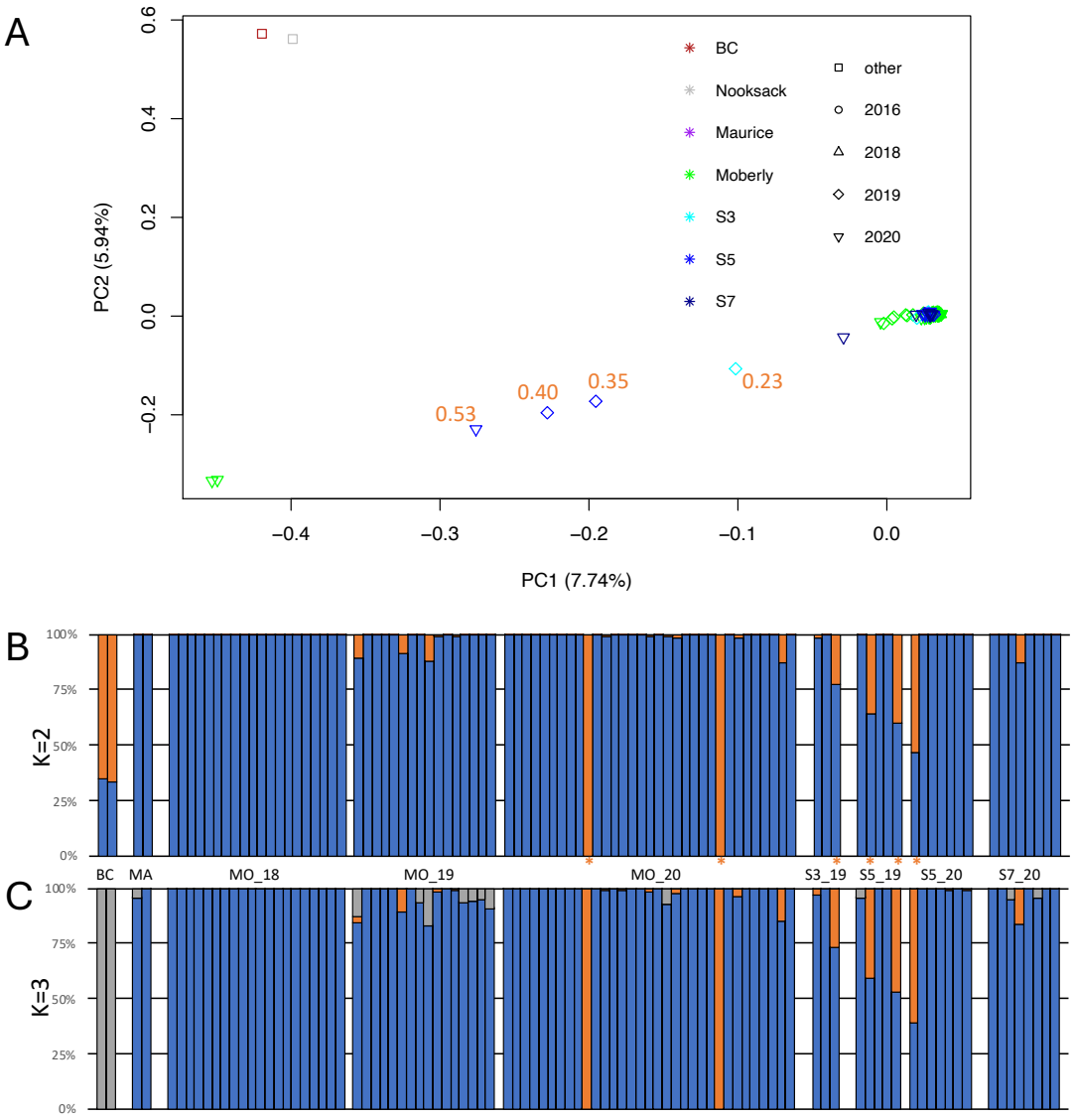

Figure 3. Population structure of Longnose Dace inferred with the SNP96 dataset (26,074 SNPs). Samples were collected in the Moberly River in 2018 (MO_18, N=20), in 2019 (MO_19, N=16) and in 2020 (M_20, N=33), in Maurice Creek in 2006 (MA, N=2), in the Peace River Section 3 in 2019 (S3_19, N=3), in Section 5 in 2019 (S5_19, N=5) and in 2020 (S5_20, N=7), and in Section 7 in 2020 (N=8). One Longnose Dace from southwestern BC (BC in PCA and Admixture plots) and one Nooksack Dace (also from southwestern BC, Nooksack in PCA and BC in Admixture plots). Panel A 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 and the sampling year by different symbols. Numbers in orange indicate ancestry proportions (Q value) from the Admixture analysis with K=2 (in panel B) for samples that appear as reticulations between the samples on the extremes of the long internal branch in Figure 2. The bottom panels show the results of an Admixture analysis with (B) two genetic groups (K=2) and with (C) three genetic groups (K=3). 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 group. Asterisks indicate the four samples whose admixture proportions are indicated in Figure 2 and Figure 3A.

To get a clearer estimate of population structure in the LAA for samples that showed little or no signs of admixture with the divergent Longnose Dace lineage identified above, we generated a new dataset with 83 samples that excluded samples from outside the LAA and samples from the LAA that had more than 2% inferred ancestry in that divergent genetic lineage (Table 7 and Appendix IV). This resulted in 50,575 SNPs before pruning for LD (used for estimates of the genetic differentiation index F_{ST}) and 22,148 unlinked SNPs for PCA and Admixture analyses.

Results from the PCA (Figure 4; Appendix IV) on the SNP83 dataset recapitulated the results along the 5th principal component of the SNP96 analysis described above, i.e., it separated along the first principal component (PC1, explaining 1.8% of variation in the data), though imperfectly, samples collected in the Moberly River from samples collected in the Peace River mainstem and Maurice Creek. Additional axes of variation did not provide additional sample clustering according to geography or sampling year.

Results from the Admixture analyses agreed with the PCA (Figure 4; Appendix IV). Though a model with a single genetic group (K=1, had the lowest CVE) was the best fit to the data, suggesting that no population structure exists among these 83 samples of Longnose Dace from the LAA, a model with two genetic groups (K=2) mirror those from PC1. In that analysis, 17 out of 20 samples collected from Maurice Creek and the mainstem of the Peace River are inferred to have all their ancestry from one genetic group (in orange in Figure 4) and 43 out of 63 samples from the Moberly River are inferred to have all their ancestry from the other genetic group (in blue in Figure 4).

Samples collected from most sampling locations were inferred to have some degree of mixed ancestry between the two groups.

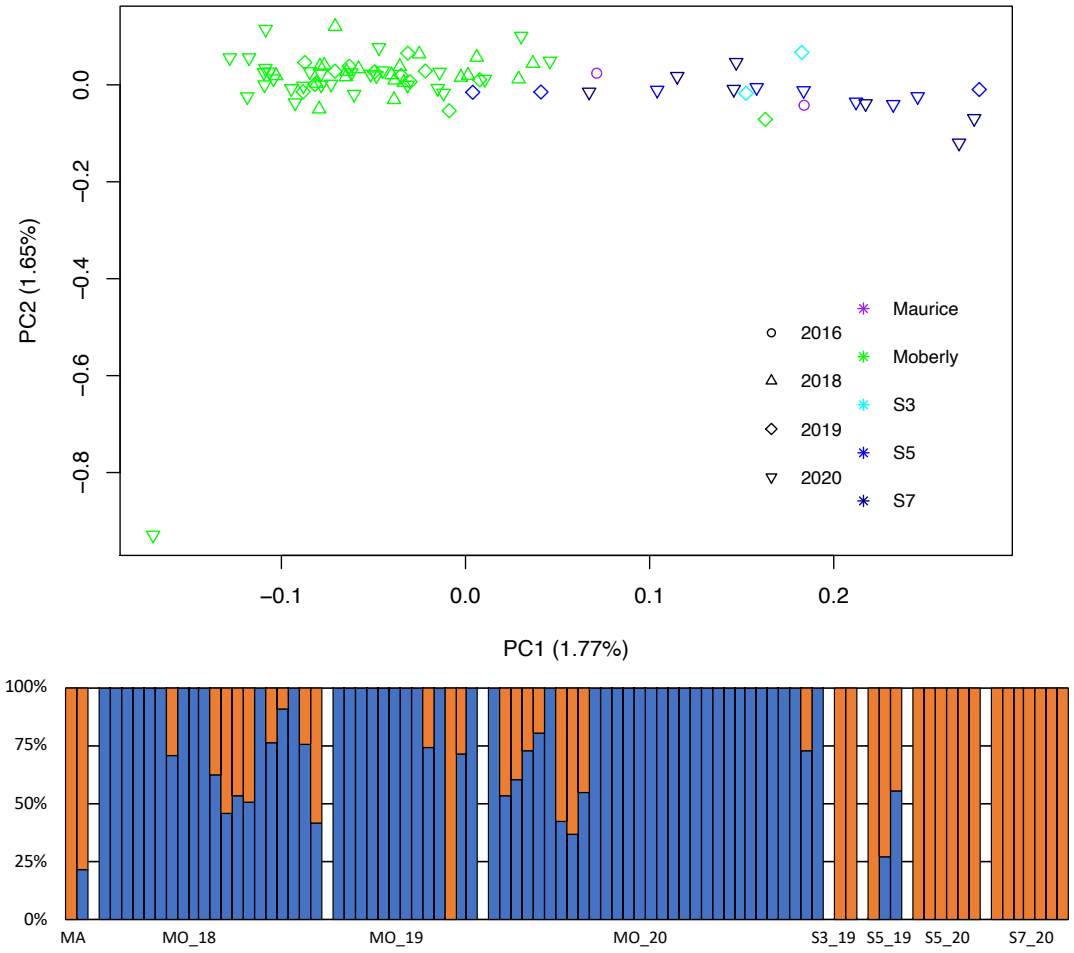

Figure 4. Population structure of Longnose Dace inferred with the SNP83 dataset (22,148 SNPs). Samples were collected in the Moberly River in 2018 (MO_18, N=20), in 2019 (MO_19, N=13) and in 2020 (M_20, N=30), in Maurice Creek in 2006 (MA, N=2), in the Peace River Section 3 in 2019 (S3_19, N=2), in Section 5 in 2019 (S5_19, N=3) and in 2020 (S5_20, N=6), and in Section 7 in 2020 (N=7). 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 and the sampling year by different symbols. The bottom panels show the results of an Admixture analysis with two genetic groups (K=2). 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 group.

Levels of genetic differentiation between sampling regions were estimated using all 50,575 SNPs in the SNP83 dataset that remained after filtering for population

genetics analysis but prior to LD pruning. Weir and Cockerham weighted F_{ST} (Weir and Cockerham, 1984) estimates (Table 8) ranged from a low of 0 between Sections 5 and 7 of the Peace River (indicating no genetic differentiation between them) to 0.015 and 0.017 between the Moberly River and Sections 5 and 7 of the Peace River. Sample sizes for the remaining sampling regions were too low (fewer than five samples available) to confidently estimate F_{ST} values. These results are in good agreement with those of the PCA and Admixture analyses analysis in showing very low levels of genetic differentiation in the LAA.

Table 8. Weighted average Weir and Cockerham's F_{ST} between sampling regions and years of Longnose Dace in the LAA estimated for the SNP83 dataset with unlinked polymorphic loci with minor allele frequency above 1%.

	Moberly	Peace S5
Peace S5	0.0148	
Peace S7	0.0174	-0.0009

ADDITIONAL WORK

Materials and Methods

As per Geraldes and Taylor (2023), we took advantage of other GBS projects in the lab to generate SNP data for three char samples (*Salvelinus* spp.) collected in Section 5 of the Peace River in 2022. The field sampling crew suggested that these were char samples (*Salvelinus* spp.), but the species could not be unambiguously identified. We used GBS following the protocols in previous reports (e.g. Geraldes and Taylor 2023) to sequence (in a sequencing project independent from this study) the samples 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 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 files for all 36 samples a) the three *Salvelinus* spp. samples collected in Section 5 of the Peace River in 2022, b) eight Lake Trout samples (*Salvelinus namaycush*; one sample from Geraldes and Taylor 2023 and seven new samples from Geraldes and Taylor, unpubl.), c) five Brook Trout (*Salvelinus fontinalis* two samples from Geraldes and Taylor 2023 and three new samples from Geraldes and Taylor, unpubl.), d) four Bull Trout samples (two from the LAA from Geraldes et al. 2020 and two from coastal BC populations from Geraldes and Taylor, unpubl.), e) 12 Dolly Varden samples (*Salvelinus malma* from three different genetic lineages/subspecies; all 12 samples from Geraldes and Taylor, unpubl.), and f) seven Arctic char (*Salvelinus*

alpinus from Geraldes and Taylor, unpubl.). Samples b) through f) served as reference samples to assist in identifying the unknown samples 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 is observed only once), 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 36 *Salvelinus* spp. samples with our filtered phylogenetic dataset (13,956 SNPs). The resulting network clearly shows each of the four species as a separate monophyletic group (Figure 5) separated from other such groups by long internal branches. All three *Salvelinus* spp. samples collected in Section 5 of the Peace River in 2022 grouped with Lake Trout samples and were thus inferred to be Lake Trout.

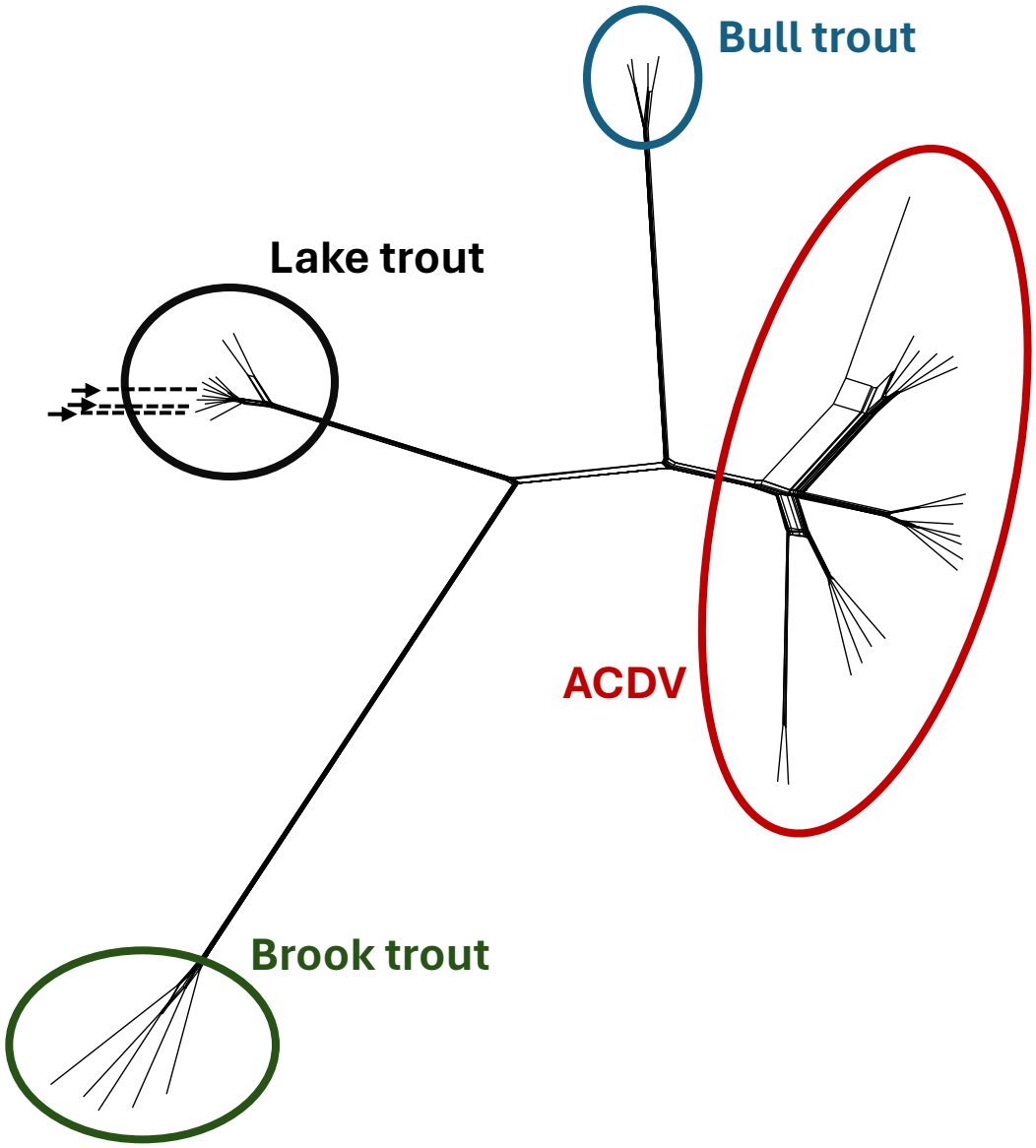

Figure 5. Phylogenetic network (13,956 SNPs) of 36 char samples showing that three samples collected in Section 5 of the Peace River in 2022 (indicated by black arrows and dotted lines) are Lake Trout. Monophyletic groups are indicated by coloured circles (black for Lake Trout, blue for Bull Trout, green for Brook Trout and dark red for Arctic Char and Dolly Varden [ACDV]).

DISCUSSION

Consistent with results from previous years, our analyses of samples collected in 2022 from various sections of the Peace River mainstem found that the vast majority of Bull Trout, Arctic Grayling, and Rainbow Trout originated from spawning tributaries upstream of the Project. The same pattern held true for the smaller number of fish of all three species collected from the TUF. Although for the first time, no Arctic Grayling were assigned to downstream of the Project, there were too few fish (only one) collected downstream of the Project to attach any significance to this finding.

Also consistent with previous years' analyses were the findings that the Halfway and Moberly rivers are key tributaries for the production of Bull Trout and Arctic Grayling, respectively, and that assignment of Rainbow Trout to UP or DP produced the highest percentage of unassigned fish (e.g., 22% for Rainbow Trout versus 3% and 0% for Bull Trout and Arctic Grayling, respectively), likely a result stemming from the stocking of hatchery strains of Rainbow Trout within the LAA and adjacent areas (see Geraldes and Taylor 2022).

Our initial results for developing GT-Seq panels for Bull Trout and Rainbow Trout are promising. A total of 189 autosomal loci and one sex-linked locus were developed for Bull Trout. Genotyping success was high and discrepancies between GT-Seq and GBS-based genotypes were low. It is important to note that GT-Seq assays require larger amounts of high-quality DNA so future sampling requirements will need to be assiduously adhered to (minimum tissue sizes, ratio of tissue to ethanol, etc). The few discrepancies between GT-Seq and GBS genotypes for the same individuals are most likely explained by uncertainty in the GBS genotype calls given the relatively low

sequence coverage obtained during our initial GBS work on Bull Trout (Geraldes and Taylor 2020). Bootsma et al. (2020) reported a concordance between GT-Seq and restriction site associated DNA sequencing (RAD, a similar technique to GBS) generated genotypes in Walleye (*Sander vitreus*) of 96.6%, very similar to our level of about 97-98%, and stated that greater coverage of initial RAD data would increase concordance further. Similarly, the determination of sex showed almost perfect concordance between field-determined and genetic sex identification (75/76 fish successfully genotyped matched between the field and genetics) and the one discrepancy could result from either method with, perhaps, a miss-sexed fish under field conditions more likely. Our next steps will be to: (i) apply the GT-Seq panel in demographic analyses in Bull Trout (estimates of effective population site and genetic diversity per locality), (ii) provide estimates of parentage assignments between adult and juvenile samples, and (iii) complete the design of a similar panel for Rainbow Trout in the LAA.

Our work further expanded this year with the analysis of the second of three non-game, small-bodied fishes in the LAA – the Longnose Dace. As discussed in Geraldes and Taylor (2023), 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 (e.g., Ruzich et al. 2019).

The Longnose Dace has been the subject of a number of phylogeographic and population genetic analyses. For instance, Bartnik (1972) studied the breeding behaviour of Longnose Dace in two populations, one from east of the Continental Divide (Manitoba) and one from west of the divide (southwestern BC). Bartick (1972) reported

differences in spawning colouration of males and breeding time (nocturnal versus diurnal) and behaviour and suggested that these differences were consistent with recognition of the two groups of Longnose Dace as two subspecies: *R. c. cataractae* for eastern dace and *R. c. dulcis* for fish found west of the divide. Although subspecies of Longnose Dace are not widely recognized in Canada, more recent investigations have indicated the presence of distinct phylogeographic lineages in Canada, especially the Longnose and Nooksack Daces in southwestern BC and adjacent portions of Washington State (Taylor et al. 2015). These authors also indicated that a major genetic distinction occurred between Longnose Dace sampled east and west of the Continental Divide. A similar result was reported in the more extensive geographic survey of Kim and Conway (2014) that recognized major “western” and “eastern” lineages who last shared a common ancestor about 2-3 million years ago (i.e., mid-Pleistocene to late Pliocene). Kim and Conway (2014) speculated that Longnose Dace originated west of the Continental Divide and that large east flowing rivers such as the Peace River were likely critical to the eastward expansion of one of the mostly widely distributed freshwater fishes in North America. In Canada, representative populations of both lineages were still quite distant from the LAA, the Columbia River basin for representative western lineage Longnose Dace and the Churchill River (Manitoba) for the eastern lineage, although they were about half again as close in the US samples (Kim and Conway 2014; Taylor et al. 2015). Given that the Peace River within the range of the LAA is known as an area of contact between fish faunas originating east and west of the Continental Divide (e.g., Rempel and Smith 1986; Foote et al. 1992; McPhail 2007), it is possible that the two major genetic groups of Longnose Dace that we

detected in the LAA are representative of the aforementioned eastern and western lineages. We will test this hypothesis with the next round of GBS results (submitted for an unrelated project) where we have included more samples of eastern Longnose Dace identified in earlier work (Ontario, Québec), including some closer to the divide (e.g., Kananskas Lake, Alberta; Taylor et al. 2015) and further samples of dace from the LAA and points further west.

On a more local geographic scale, population structure of Longnose Dace has been investigated both at presumptively neutral loci (microsatellites) and at adaptive loci (major histocompatibility complex, MHC) with somewhat conflicting results. For instance, Crispo et al. (2017) examined population structure of Longnose Dace within the Bow, Oldman (both of the South Saskatchewan River [Hudson Bay] drainage) and Peace rivers in Alberta. These authors reported no isolation-by-distance (IBD) within either the Oldman or Bow rivers (Peace River sites were inadequate to test for IBD) across distances of about 150 km (although the Oldman River relationship was modestly positive at 0.34 and approached significance, $P = 0.07$). Further, a model-based analysis similar to our Admixture analyses reported that structure was only evident at a level of that between the Peace River and all other samples (including both the North and South Saskatchewan rivers), but similar analyses *within* each of these basins were not attempted (Crispo et al. 2017). By contrast, the authors reported significant allelic variation at MHC loci at sites within drainages upstream and downstream of municipal and agricultural areas associated with different levels of contamination (see also Girard and Angers 2011). As such, the authors detected no, to

weak, population structure at neutral loci, but clear structure at loci possibly subject to natural selection through the contaminant environmental gradient.

Weiman and Berendzen (2018) reported significant IBD (r = 0.37) and population subdivision (F_{ST} = 0.012 to 0.043), and the presence of two major genetic groups (via admixture analysis) by assaying microsatellites in Longnose Dace collected from 15 localities spanning tributaries of the upper Mississippi River in the “Driftless Area” in northwestern Iowa, a region bypassed during the most recent glacial advance in North America. The tributaries spanned about 140 km of the upper Mississippi River and the authors concluded that the resultant genetic structure was primarily a function of historical processes (glacial history and colonization history) rather than contemporary processes or human-induced habitat changes. Finally, Ruskey and Taylor (2016) examined microsatellite DNA variation among populations of Nooksack and Longnose Dace in BC and Washington State. Among Longnose Dace samples located within interconnected tributaries of the lower Fraser River, significant population structure was evident (e.g., F_{ST} reached up to 0.08 between tributaries separated by 65 km of mainstem Fraser River and were as low as 0.015 between tributaries separated by 8.5 km of mainstem Fraser River).

These comparative studies have several implications for our results. First, the lack of apparent structure amongst localities within the mainstem Peace River is consistent with that observed among sites within the Oldman and Bow rivers in southwestern Alberta (Crispo et al. 2017). Second, that we detected significant structure between the tributary Moberly River and the mainstem Peace River is consistent with studies in other areas (Iowa, Weiman and Berendzen (2018); lower Fraser River,

Ruskey and Taylor (2016)). Third, as shown in the behavioural studies of Bartnik (1972) and the genetic studies of Kim and Conway (2014), Taylor et al. (2015), and Weiman and Berendzen (2018), historical processes may be important in driving the observed genetic structure of Longnose Dace, including that in the LAA. Finally, the work of Crispo et al. (2017; see also Girard and Angers 2011) highlight the fact that the lack of population genetic structure at neutral loci does not mean that significant genetic structure does not exist at loci subject to natural selection (i.e., MHC loci in the cited studies). Furthermore, the distinction we observed between Longnose Dace sampled from the Moberly and Peace rivers mirrors, in a relative sense, that for Slimy Sculpin although the absolute difference between the two localities was an order of magnitude lower for Longnose Dace (Geraldes and Taylor 2023). The concordance in these findings between species suggests that they may reveal general deterministic phenomena that drive the evolution of similar patterns in distinct species (see Zbinden et al. 2023) whether they be primarily demographic (e.g., resistance to downstream migration from the Moberly to the Peace or upstream from the Peace to the Moberly in characteristically benthically-oriented species) or a common response to some aspect of natural selection acting via environmental differences between the two rivers. In this regard, the patterns revealed in a more mid-water oriented and presumably more mobile species, the Redside Shiner in our work to come, may be particularly informative. Regardless of its ultimate cause, and as we emphasized in Geraldes and Taylor (2023) for Slimy Sculpin, the differentiation between Longnose Dace 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 continues to provide 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 Longnose Dace and provided additional tools and results for *Salvelinus* species identification in the LAA. In the coming months, assignments will continue for samples collected in 2023, and we will be analyzing the GT-Seq panel genotyping results for Bull Trout and continue work we have started on: (i) developing similar GT-Seq assays for Rainbow Trout to examine demographic characteristics (e.g., effective population size, genetic variation, parentage), and (ii) population structure of the Redside Shiner.

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

Bartnik VG (1972) Comparison of the breeding habits of two subspecies of longnose dace, *Rhinichthys cataractae*. Canadian Journal of Zoology 50:83-86. https://doi.org/10.1139/z72-015

Bohling J, Von Bargen J, Piteo M, Louden A, Small M, Delomas TA, Kovach, R (2021) Developing a standardized single nucleotide polymorphism panel for rangewide genetic monitoring of bull trout. North American Journal of Fisheries Management 41: 1920-1931. https://doi:10.1002/nafm.10708

Bootsma ML, Gruenthal KM, McKinney GJ, Simmons L, Miller L, Sass GG, Larson, WA (2020) A GT-seq panel for walleye (*Sander vitreus*) provides important insights for efficient development and implementation of amplicon panels in non-model organisms. Molecular Ecology Resources 20:1706-1722. https://doi.org/10.1111/1755-0998.13226

Campbell NR, Harmon SA, Narum SR (2014) Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources 15:855–867. https://doi.org/10.1111/1755-0998.12357

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

Chang SL, Ward HGM, Russello MA (2021) Genotyping-in-Thousands by sequencing panel development and application to inform kokanee salmon (*Oncorhynchus nerka*) fisheries management at multiple scales. PLoS One 16: e0261966 (2021). https://doi.org/10.1371/journal.pone.0261966

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

Crispo E, Tunna HR, Hussain N, Rodriguez SS, Pavey SA, Jackson LJ, Rogers SM (2017) The evolution of the major histocompatibility complex in upstream versus downstream river populations of the longnose dace. Ecology and Evolution 7: 3297-3311. https://doi.org/10.1002/ece3.2839

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

Evans, M.L., Johnson, M.A., Jacobson, D., Wang, J., Hogansen, M. and O'Malley, K.G., 2016. Evaluating a multi-generational reintroduction program for threatened salmon using genetic parentage analysis. Can J Fish Aquat Sci 73:.844-852. https://doi.org/10.1139/cjfas-2015-0317

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

Foote CJ, Clayton JW, Lindsey CC, Bodaly RA (1992) Evolution of lake whitefish (*Coregonus clupeaformis*) in North America during the Pleistocene: evidence for a Nahanni glacial refuge race in the northern Cordillera region. Canadian Journal of Fisheries and Aquatic Sciences 49:760-768. https://doi.org/10.1139/f92-085

Geraldes A, Taylor EB (2020) Fish genetics study 2019 annual report. British Columbia Hydro Site C Clean Energy Report Series. https://www.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://www.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. https://sitecproject.com/sites/default/files/fish-genetics-study-2021-annual-report.pdf. Accessed April 20, 2023.

Geraldes A, Taylor EB (2023) Fish genetics study 2022 annual report. British Columbia Hydro Site C Clean Energy Report Series. https://www.sitecproject.com/sites/default/files/Fish-Genetics-Study-2022-Annual-Report.pdf. Accessed April 20, 2023.

Girard P, Angers B (2011) The functional gene diversity in natural populations over postglacial areas: the shaping mechanisms behind genetic composition of longnose dace (*Rhinichthys cataractae)* in northeastern North America. Journal of Molecular Evolution73:45-57. https://doi.org/10.1007/s00239-011-9456-1

Horn, RL, Nuetzel HM, Johnson B, Kamphaus C, Lovrak J, Mott K. Newsome T, Narum SR (2024) Utility of parentage-based tagging for monitoring Coho salmon (*Oncorhynchus kisutch*) in the interior Columbia River basin. Evolutionary Applications 17:p.e13607. https://doi.org/10.1111/eva.13607

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

Hohenlohe PA, Funk WC, and Rajora OP (2021) Population genomics for wildlife conservation and management. Molecular Ecology 30:62-82. https://doi.org/10.1111/mec.15720

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

Kim D, and Conway KW (2014) Phylogeography of *Rhinichthys cataractae* (Teleostei: Cyprinidae): preglacial colonization across the Continental Divide and Pleistocene diversification within the Rio Grande drainage. Biological Journal of the Linnean Society 111:317–333. https://doi.org/10.1111/bij.12209

Liu, A., Geraldes, A. & Taylor, E. B. Development and testing of diagnostic molecular markers to differentiate among three species of *Salvelinus*: Arctic char, Dolly Varden, and Bull trout. Conservation Genetic Resources 15: 83–96 (2023). https://doi.org/10.1007/s12686-023-01305-7

May S A, Shedd KR, Rand PS, and Westley PAH (2023) Tidal gradients, fine-scale homing and a potential cryptic ecotype of wild spawning pink salmon (*Oncorhynchus gorbuscha*). Molecular Ecology 32: 5838–5848. https://doi.org/10.1111/mec.17154

McPhail JD (2007) *The freshwater fishes of British Columbia*. Edmonton: University of Alberta Press.

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

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

Rempel LL and Smith DG (1998) Postglacial fish dispersal from the Mississippi refuge to the Mackenzie River basin. Canadian Journal of Fisheries and Aquatic Sciences 55:893-899. https://doi.org/10.1139/f97-257

Rieman BE, Allendorf FW (2001) Effective population size and genetic conservation criteria for bull trout. North American Journal of Fisheries Management 21:756-764. https://doi.org/10.1577/1548-8675(2001)021<0756:EPSAGC>2.0.CO;2

Ruskey JA, Taylor EB (2016) Morphological and genetic analysis of sympatric dace within the *Rhinichthys cataractae* species complex: a case of isolation lost. Biological Journal of the Linnean Society 117:547-563. https://doi.org/10.1111/bij.12657

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

Shedd, K. R. et al. Reduced relative fitness in hatchery-origin Pink Salmon in two streams in Prince William Sound, Alaska. Evol. Appl. 15, 429–446 (2022). https://doi.org/10.1111/eva.13356

Taylor EB, McPhail JD, Ruskey JA (2015) Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America—the Nooksack dace problem. Canadian Journal of Zoology 93:727-734. https://doi.org/10.1139/cjz-2015-0014

Weigel D, Koch I, Monzy F, Sharpe C, Narum S, Caudill CC (2019) Evaluation of a trap-and-transport program for a threatened population of steelhead (*Oncorhynchus mykiss).* Conservation Genetics 20:1195-1199. https://doi.org/10.1007/s10592-019-01200-5

Wieman AC, Berendzen PB (2018) Spatial genetic variation and habitat association of *Rhinichthys cataractae*, the longnose dace, in the Driftless Area of the upper Mississippi River basin. Conservation Genetics 19: 1367-1378. https://doi.org/10.1007/s10592-018-1106-3

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

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. https://doi.org/10.1093/bioinformatics/bts606

Zbinden ZD, Douglas MR, Chafin TK, Douglas ME (2023) Riverscape community genomics: A comparative analytical approach to identify common drivers of spatial structure. Molecular Ecology 32: 6743-6765. https://doi.org/10.1111/mec.16806