



Whole-Genome Resequencing to Evaluate Life History Variation in Anadromous Migration of *Oncorhynchus mykiss*

Erin E. Collins^{1*}, Nicolas Romero², Joseph S. Zendt² and Shawn R. Narum¹

¹Hagerman Genetics Laboratory, Columbia River Inter-Tribal Fish Commission, Hagerman, ID, United States, ²Yakama Nation Fisheries, Yakima/Klickitat Fisheries Project, Klickitat, WA, United States

OPEN ACCESS

Edited by:

Narongrit Muangmai,
Kasetsart University, Thailand

Reviewed by:

Timothy D. Leeds,
United States Department of
Agriculture (USDA), United States
Matt Hale,
Texas Christian University,
United States

*Correspondence:

Erin E. Collins
ecollins@critfc.org

Specialty section:

This article was submitted to
Evolutionary and Population Genetics,
a section of the journal
Frontiers in Genetics

Received: 15 October 2021

Accepted: 24 January 2022

Published: 15 March 2022

Citation:

Collins EE, Romero N, Zendt JS and
Narum SR (2022) Whole-Genome
Resequencing to Evaluate Life History
Variation in Anadromous Migration of
Oncorhynchus mykiss.
Front. Genet. 13:795850.
doi: 10.3389/fgene.2022.795850

Anadromous fish experience physiological modifications necessary to migrate between vastly different freshwater and marine environments, but some species such as *Oncorhynchus mykiss* demonstrate variation in life history strategies with some individuals remaining exclusively resident in freshwater, whereas others undergo anadromous migration. Because there is limited understanding of genes involved in this life history variation across populations of this species, we evaluated the genomic difference between known anadromous ($n = 39$) and resident ($n = 78$) *Oncorhynchus mykiss* collected from the Klickitat River, WA, USA, with whole-genome resequencing methods. Sequencing of these collections yielded 5.64 million single-nucleotide polymorphisms that were tested for significant differences between resident and anadromous groups along with previously identified candidate gene regions. Although a few regions of the genome were marginally significant, there was one region on chromosome Omy12 that provided the most consistent signal of association with anadromy near two annotated genes in the reference assembly: COP9 signalosome complex subunit 6 (CSN6) and NACHT, LRR, and PYD domain-containing protein 3 (NLRP3). Previously identified candidate genes for anadromy within the inversion region of chromosome Omy05 in coastal steelhead and rainbow trout were not informative for this population as shown in previous studies. Results indicate that the significant region on chromosome Omy12 may represent a minor effect gene for male anadromy and suggests that this life history variation in *Oncorhynchus mykiss* is more strongly driven by other mechanisms related to environmental rearing such as epigenetic modification, gene expression, and phenotypic plasticity. Further studies into regulatory mechanisms of this trait are needed to understand drivers of anadromy in populations of this protected species.

Keywords: *Oncorhynchus*, steelhead, anadromy, whole-genome resequencing, Klickitat River

1 INTRODUCTION

Variation in life history traits is often advantageous to the persistence of natural species, and salmonids retain diverse life history types to contend with environmental stochasticity (e.g., Moore et al., 2014). Species such as *Oncorhynchus mykiss* demonstrate variation in migratory life history with some individuals remaining exclusively resident in freshwater, whereas others exhibit anadromous migration from freshwater rearing environments to the ocean before returning to natal areas to spawn. Populations can consist of primarily anadromous or resident fish, or there can be a mixture of resident and anadromous individuals within a single population. Anadromy is rare, <1% of all fish species are anadromous (McDowall 1988), because the biological cost includes challenges of surviving in both fresh and saltwater habitats (e.g., physiological changes to regulate ions, long distance migrations, and predation rates) and must be outweighed by benefits such as improved productivity and access to seasonally available resources to feed and spawn. Because of smoltification and migration of anadromous fish, reproduction is delayed for adults; however, they benefit from large size and high fecundity gained from highly productive marine feeding habitat. Ultimately, life history development is a complex trait and is influenced by genetics, environment, and individual fitness (Fleming and Reynolds 2004). Variation in the degree of anadromy or residency of *O. mykiss* populations impacts the abundance, diversity, resilience, structure, and productivity of *O. mykiss* populations (Waples et al., 2008). There are greater benefits for female *O. mykiss* to migrate and grow to a larger size before spawning compared to males. Larger female *O. mykiss* experience increased fecundity, increased egg size, improved redd site protection, can dig deeper into the substrate, and can attract more mates (Quinn et al., 2011). Whereas, male *O. mykiss* are more likely than females to be resident due to an option for males to spawn at a small size with the “sneaky/sneaker” male tactic (Fleming and Reynolds 2004), and this variation in males was the focus of this study.

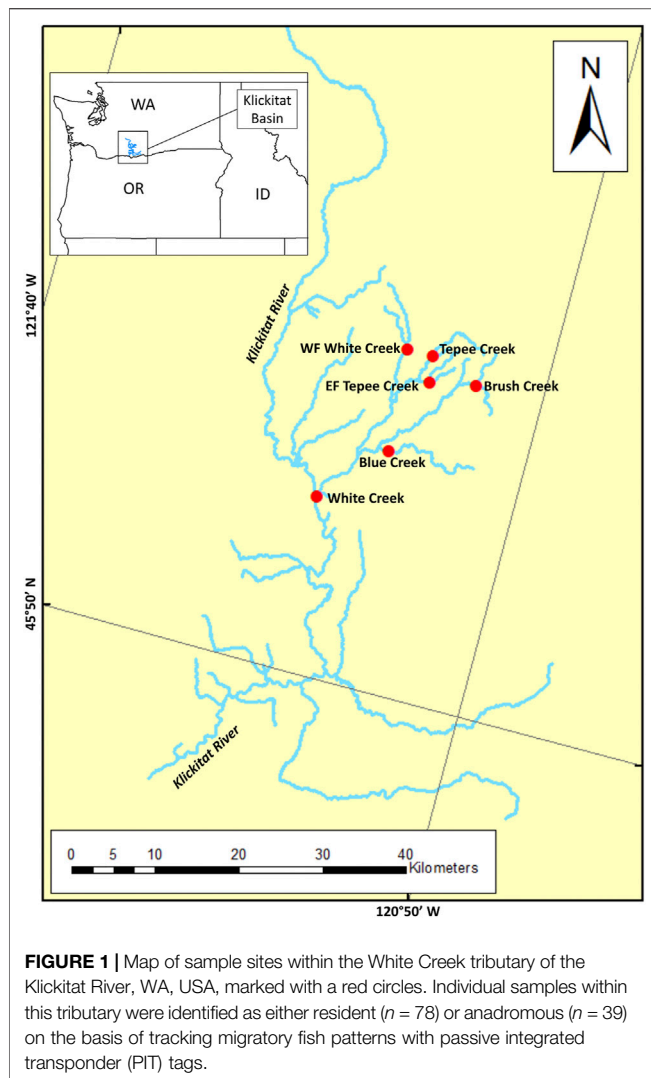
Oncorhynchus mykiss are of immense economic, ecological, and cultural importance throughout their range. Populations of *O. mykiss* are considered partially migratory due to the presence of both anadromous, “steelhead,” and resident, “rainbow trout,” forms of the species. These life history types of *O. mykiss* are sympatric and can produce offspring of either life history type (Christie et al., 2011; Courter et al., 2013; Sloat and Reeves 2014). Anadromy of *O. mykiss* are initiated by a combination of environmental and genetic cues and the genetic component is currently not well understood across much of the geographic range of this species. A continued presence of both resident and anadromous *O. mykiss* is critical to population resilience to environmental or anthropogenic disturbances (Moore et al., 2014).

Previous studies have identified candidate genes or regions in the genome associated with the trait for anadromy in steelhead, but association has been limited to particular populations. For example, Pearse et al. (2019) and Martínez et al. (2011) detected a “migration supergene” in an inversion on chromosome Omy05, where the ancestral form is associated with anadromous steelhead

and the derived form is associated with resident rainbow trout. The signal from the chromosome Omy05 inversion was significant in coastal populations from southern portions of the range but had low variability for other populations of steelhead (inland or northern latitudes). Studies of anadromy in specific populations of *O. mykiss* have identified putative candidate genes associated with resident versus migratory life history (e.g., Nichols et al., 2008; Narum et al., 2011; Hale et al., 2013), but these candidates have not been directly validated in further studies. In addition, previous studies (Nichols et al., 2008; Narum et al., 2011; Hale et al., 2013; Hecht et al., 2013) detected outlier loci of small effect throughout the genome for anadromy, indicating an extremely polygenic trait, and very few of the same loci were detected in multiple analyses and little overlap was observed between populations tested, indicating a lack of a conserved gene region or high environmental plasticity for the trait. Thus, studies to date indicate that the genetic basis of anadromy is controlled in a population specific manner. However, there is currently a paucity of data from many populations, limiting our understanding of the relative importance of shared alleles and population specific anadromy.

Steelhead in the Klickitat River, WA, USA, are geographically and genetically intermediate between coastal and inland genetic lineages (Blankenship et al., 2011; Collins et al., 2020). The geography along the Klickitat River and its tributaries is diverse and includes waterfalls that act as partial barriers to migration, leading to reproductive and genetic isolation for populations in some areas (Narum et al., 2006; Narum et al., 2008). Another genetic influence on Klickitat River *O. mykiss* populations has been the stocking of summer-run hatchery steelhead, known as Skamania steelhead, since 1961, and the Skamania steelhead may have overlapping spawning timing with natural origin summer steelhead, but with limited gene flow (Narum et al., 2006). Skamania steelhead were originally derived from steelhead from the Washougal and Klickitat Rivers and have been strongly selected for early spawning times (Crawford 1979). Resident *O. mykiss* occur throughout substantial portions of this drainage, especially in headwater tributaries with lower flows and water temperature than large tributaries or the mainstem Klickitat River (Narum et al., 2008; Evans et al., 2021). Because Klickitat River anadromous steelhead are listed as threatened and are protected under the Endangered Species Act, identifying that the genetic component of anadromy may be a useful conservation tool to assist with recovery of steelhead in this system.

In this study, we aimed to use whole-genome resequencing to assess underlying genetic mechanisms for the trait of male anadromy in *O. mykiss* from the Klickitat River. This study combined high marker density throughout the genome and phenotypic data from long-term tracking of individuals to test for regions of the genome associated with resident versus anadromous migration patterns. Further, markers from previously identified candidate genes were tested for differences among collections. A genetic sex marker on the basis of the sexually dimorphic on the Y chromosome region (sdY; Yano et al., 2013) was used to identify males and females, enabling analyses to focus on sex specific migration patterns in



males to avoid sex bias when comparing anadromy life traits due to the biological constraint of rare resident female rainbow trout.

2 MATERIALS AND METHODS

2.1 Sample Collection and Phenotypes

Samples of juvenile *O. mykiss* were collected from the Klickitat River *via* electrofishing from the following tributaries of the White Creek drainage: White Creek, West Fork White Creek, Blue Creek, Brush Creek, Tepee Creek, and East Fork Tepee Creek (Supplementary Table S1 and Figure 1). All individual *O. mykiss* were captured as juveniles (parr appearance) in multiple sites of the White Creek drainage between 2009 and 2019 and were inserted with a 12.5-mm full-duplex passive integrated transponder (FDX-B PIT) tag (Biomark, Boise, ID, USA) to track individual migration patterns and a non-lethal fin clip was collected for genetic analyses. Fish were tracked for multiples years to determine migratory classification as either resident or anadromous for genomic analyses.

Sampling was ongoing and conducted every year with marked fish continuously monitored *via* passive mark recapture methods. Captured fish were initially scanned with a handheld PIT tag scanner to identify previously PIT-tagged fish (recaptures) from non-tagged fish at each sample site. A PIT tag was inserted in each non-tagged *O. mykiss* ≥ 65 mm. Sample date, PIT tag code, and biometric data were recorded from each fish. In addition, a non-lethal fin clip for genetic analysis was collected from first time captured fish. Each sampled fish was returned to the stream at their capture location. Movement patterns were tracked through a combination of active and passive mark-recapture methods. Electrofishing surveys were conducted at each site annually to actively monitor fish. Electrofishing was the method used to both capture fish for marking and to manually recapture previously tagged fish. We used electrofishing as a method to classify resident fish by establishing a temporal record of recapture events for a given individual.

A network of PIT tag interrogation arrays in the Klickitat River and Columbia River were used to passively monitor migratory behavior. PIT tag interrogation arrays located in tributaries to the Klickitat River were designed to interrogate the entire wetted channel width *via* channel spanning PIT tag arrays. The mainstem Klickitat River floating array interrogated approximately two-thirds of the mainstem wetted channel. Sites varied from having a minimum of two arrays to a maximum of three arrays. Depending on stream width, arrays consisted of a single antenna to up to three adjacent antennas. Arrays were configured by placing antenna arrays perpendicular to the stream channel. PIT tag arrays were spaced five to 10 m apart along the longitudinal stream axis. For each fish detection, the PIT tag code, date/time stamp, and unique antenna number were recorded to a data logger buffer. Directionality of movement and temporal movement patterns were subsequently determined from these data. Juvenile bypass interrogation arrays at Bonneville Dam operate continuously, except for a short non-operational period in the winter. In addition, an array is towed through the Columbia River estuary to detect PIT-tagged fish throughout the year. All PIT tag interrogation sites were operational during the juvenile out-migration period.

Individual *O. mykiss* samples were classified as either resident or anadromous based on multiple years of mark-recapture sampling and monitoring. Anadromous fish were identified as those that successfully outmigrated from the Klickitat River and were either last detected in the mainstem Columbia River presumably headed to the ocean or returning from the ocean. Fish were classified as freshwater residents if they remained in freshwater for a minimum of 2 years between initial and final recapture events and did not register at any of the downstream PIT tag arrays.

2.2 Molecular Methods

Prior to whole-genome resequencing, specimens were genotyped with a panel of single-nucleotide polymorphism (SNP) markers to identify sex of samples and to account for kinship and population structure. A Chelex 100 method (Sigma-Aldrich, St. Louis, MO) was followed to extract DNA from the tissue of a total of 198 *O. mykiss* individuals (Supplementary Table S1). The genotyping-in-thousands by sequencing method (GT-seq; Campbell et al., 2015) was employed for all samples. Samples

were genotyped with a panel of 376 SNPs (Collins et al., 2020) that were a mix of putatively neutral and adaptive markers, a sex marker, and markers that identify closely related species of cutthroat trout (*O. clarkii*). Briefly, our study followed standard GT-seq methods that implemented two rounds of polymerase chain reaction (PCR) to first amplify targeted SNPs and then add dual barcodes for the identification of all individuals in downstream analyses. Following the barcoding step, the concentration of each sample was normalized and then pooled into a library. Once the library was prepared, it was sequenced on an Illumina NextSeq 500 instrument and then was genotyped with scripts from Campbell et al. (2015). As a quality control step, all samples and loci with $\geq 10\%$ missing genotypes were removed from further analyses. Genotype data were retained when $>90\%$ loci successfully genotyped and had an estimated $<0.5\%$ genotyping error based on replicate genotyping of 10% of samples.

To select the final samples for the whole-genome resequencing step, samples were first analyzed and filtered according to biases from kinship, population structure, and sex imbalances among samples. The sex of individuals was determined with the *O. mykiss* sex marker (OmyY1_2SEXY; Brunelli et al., 2008) with the intent to split individuals into four groups (resident male, anadromous male, resident female, and anadromous female) for statistical analysis. This approach was intended to account for sex specific differences in migration behavior and isolate confounding signals of sex-linked genes including the known sdY region on chromosome Omy29 (Gao et al., 2021). Kinship and population biases were tested at neutral loci with a compressed linear model (Zhang et al., 2010) that used a kinship matrix produced with the VanRaden algorithm (VanRaden et al., 2008) along with three principal components and were implemented in Genome Association and Prediction Integrated Tool (GAPIT-R package; Lipka et al., 2012).

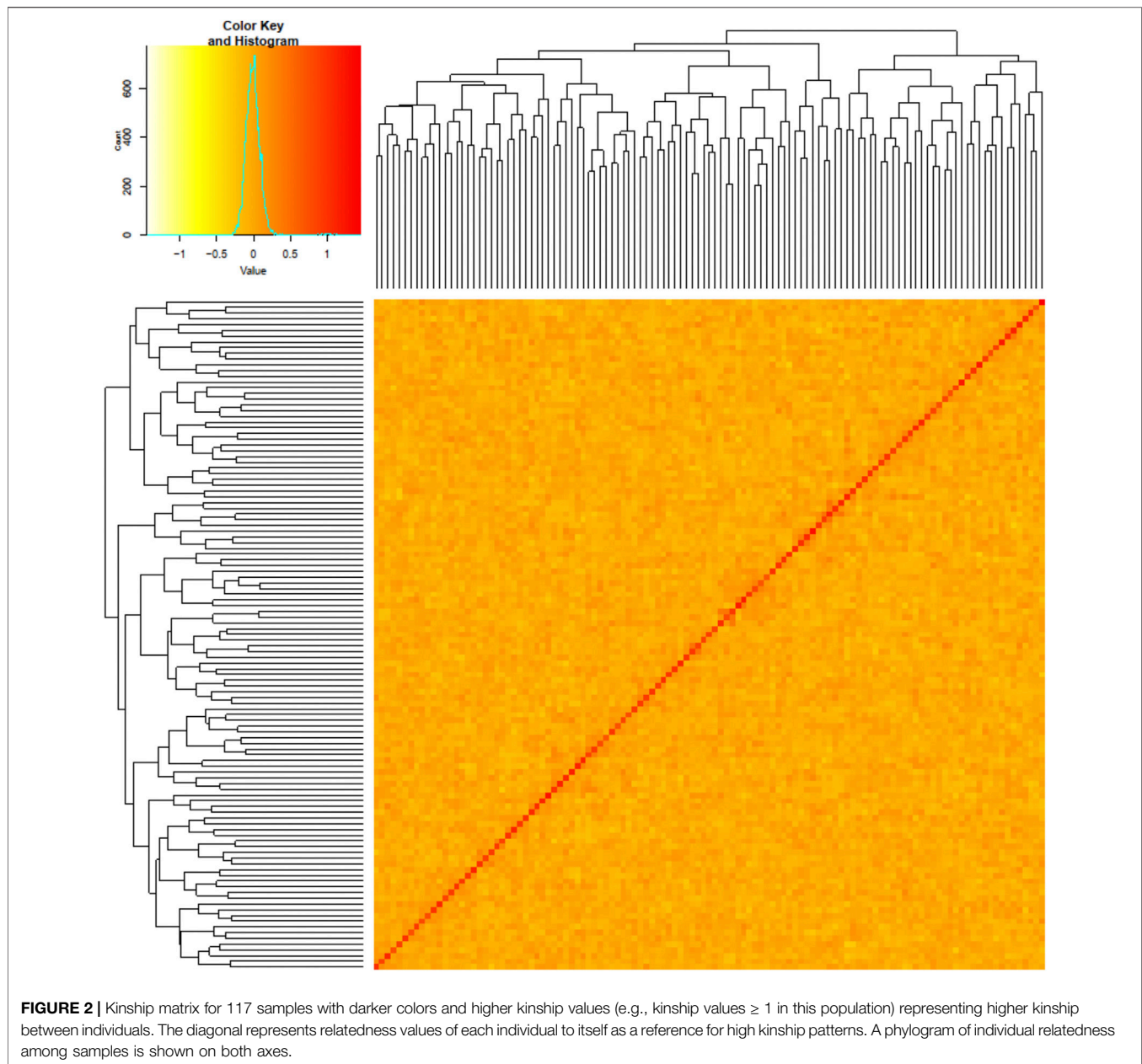
Our approach to genome resequencing was to target low read depth for individually barcoded samples and then to combine sequence reads from individuals of the same phenotype for further analyses of the four groups: resident male, anadromous male, resident female, and anadromous female. To prepare the libraries for whole-genome resequencing, individuals were barcoded, and a NEBNext Ultra enzymatic fragmentation protocol was followed (Horn et al., 2020). Briefly, we extracted DNA according to a Chelex extraction method (Sweet et al., 1996). Individual sample DNA concentrations were quantified with pico-green fluorescence on a Tecan M200 (Männedorf, Switzerland) to normalize the quantity of DNA from each sample within two standard deviations of the mean concentration. After barcoding individuals, the DNA samples were combined to avoid batch effects and normalized DNA was fragmented with NEBNext double-stranded DNA fragmentase, which fragments randomly throughout the genome. Fragments of 400–600 base pairs (bp) were size-selected, and then, ends of DNA strands were repaired with NEBNext end prep. NEBNext adaptors were ligated to the repaired and size-selected fragments and cleaned with SpriSelect Beads. We implemented PCR amplification and cleaned the PCR product with SpriSelect

Beads. The quantity and quality of each library were assessed using quantitative PCR and a Tape Station with a DNA High Sensitivity D1000 kit and then normalized prior to sequencing. Final libraries were sequenced on an Illumina NextSeq 500 sequencer, targeting 500 million paired-end reads per library.

2.3 Statistical Analyses

Barcodes were useful to normalize sequence reads for individual samples, but samples were subsequently pooled for analyses by migratory traits to provide adequate sequencing depth to identify SNPs and estimate allele frequencies for each phenotypic group. The whole-genome resequencing data were analyzed with a pipeline, called PoolParty (Micheletti and Narum 2018), that can handle either pooled or individually barcoded sequence data. The PoolParty pipeline consists of bash scripts that utilize open-source packages to align to a reference genome and filter raw whole-genome resequencing data. In this study, sequence reads were aligned to *O. mykiss* reference assembly GCA_013265735.3 (Gao et al., 2021). Next, the pipeline assesses filtered alignments with BMap (Bushnell 2016), bwa mem (Li 2013), samblaster (Faust and Hall 2014), Picard Tools (Broad Institute), and bcftools (Li et al., 2011). Mapped-read coverage statistics were determined to evaluate mapping efficiency to the reference genome with mean depth of coverage and proportion of the reference assembly genome covered by mapped reads and filtered accordingly. To filter the SNPs that were used in downstream analyses, we assessed coverage and minor allele frequency (MAF). When SNPs had coverage either below 15X or above 100X, or MAF below 0.05, they were filtered out of analyses. Pairwise analyses, such as the fixation index (F_{ST}), sliding window F_{ST} (sFST), Fisher's exact test (FET), and local score (Fariello et al., 2017), were conducted for resident versus anadromous groups with Popoolation2 scripts (Kofler et al., 2011). To determine significance of difference between SNPs in resident and anadromous groups, we used false discovery rate (FDR) corrected p -values from FET analyses of allele frequencies for individual SNPs. A QQ plot analysis was conducted to evaluate the potential bias of association tests and effects of covariates.

To test for polygenic signals of association with anadromy, we analyzed the data set with an R package called Random Forest (RF; Liaw and Wiener 2002). This RF package is a machine learning algorithm that provides a nonparametric framework to identify epistasis through identification of a group of loci predictive of phenotypic traits with significant differences in allele frequencies that better predict the trait than a single locus (Brieuc et al., 2018). To detect these loci, a “forest” of classification or regression trees is partitioned recursively by randomly sampling, with replacement, creating a training subset of data, and a random selection of a subset of predictors to find which partition of loci best predicts the trait by minimizing within-group variance or error. The RF algorithm first identifies the best predictive loci partition as the first node of the “tree” and then continues to randomly subset the remaining predictors. Selections of partitions were random to avoid bias and curtail differences within and between the “trees” (Goldstein et al., 2011; Rokach 2016). An error rate, a proportion of the variation explained (PVE), and an importance value were



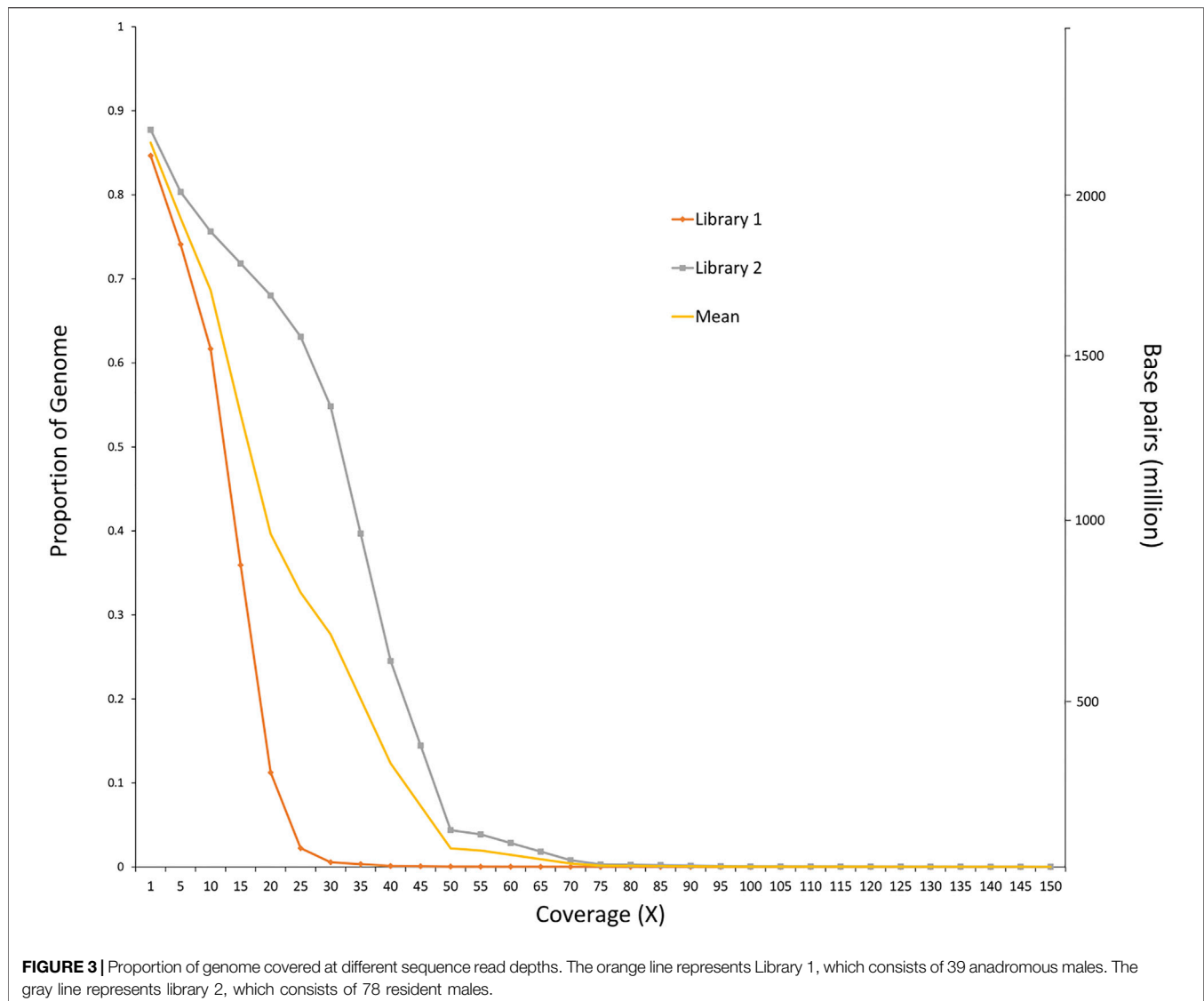
calculated with a subset of samples left out of the “tree” building steps as training data. The downside to random selections is that, if an important predictor is removed, then the predictive power of the “tree” will decrease; this can be belied with analyses of many “trees” and importance values of partitions across the “forest,” instead of relying too heavily upon a single “tree”.

To examine previously identified candidate genes for anadromy, we analyzed genotypes from candidate markers that were included in the GT-seq panel across phenotypic groups. This included several markers from the genome region described by Pearse et al. (2019) for large inversions on chromosome Omy05 of coastal California *O. mykiss* populations associated with the anadromy phenotype. In addition, we tested the three candidate loci from Narum et al.

(2011) that had previously shown significant association with anadromy in previous analyses in *O. mykiss* from the Klickitat River. For these candidate markers that were genotyped with GT-seq, we assessed the MAF across all individuals and tested for differences in allele and genotype frequencies for anadromous and resident groups with FDR correction for multiple tests.

3 RESULTS

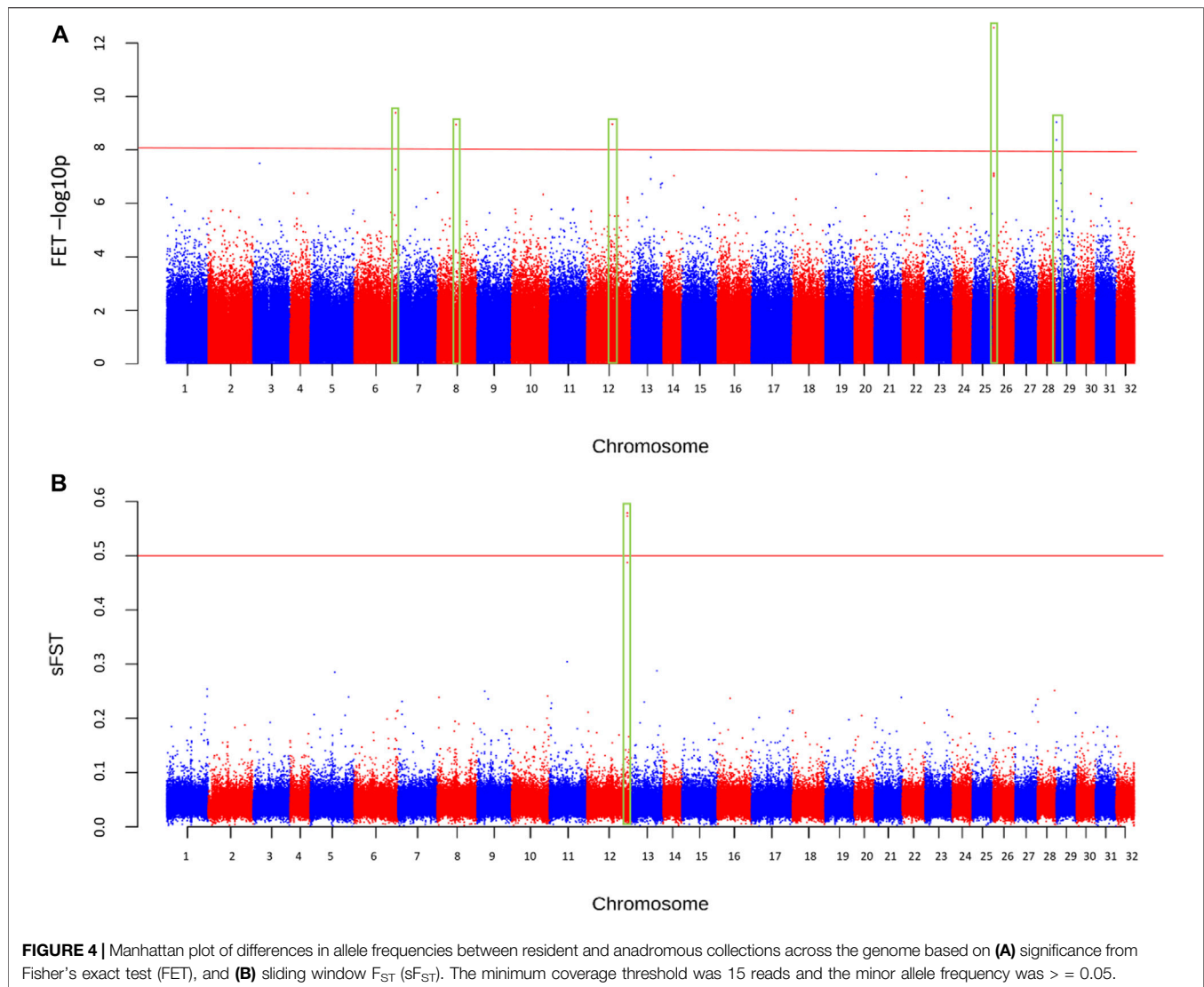
Results from genotyping the initial GT-seq panel were used to select the individuals for whole-genome resequencing. For the samples genotyped with GT-seq, the average genotype error rate was 0.51% and individuals missing $\geq 10\%$ genotypes were



excluded from analyses. Of the 198 individuals genotyped, three failed, leaving 195 for further analyses. Of those 195 individuals, the marker for genetic sex identified 117 males (78 resident and 39 anadromous) and 78 females (3 resident and 75 anadromous). This result reflected lower variation in migration life history for females than males (**Supplementary Table S1**) and a limited sample size of resident females that excluded intended analyses within the female groups (resident vs. anadromous). Analyses for potential kinship bias or population structure revealed minimal signal among the samples (**Figure 2**), and no individuals were excluded from further analyses.

The final sample set included 117 male *O. mykiss* that were split into two phenotypic groups with 39 anadromous (4 returning to spawn and 35 out-migrating) steelhead and 78 resident rainbow trout. Because there were insufficient resident females ($n = 3$) to test for migratory differences within females (resident vs. anadromous), an attempt was made to combine

females with males in their respective phenotypic group even though this caused a severe imbalance in sex ratio between groups of residents (3 female, 78 male; 3.7% female) versus anadromous (75 female, 39 male; 67.8% female). Imbalanced sex ratios in analyses were expected to lead to detection of sex-linked regions as demonstrated in previous studies (Benestan et al., 2017) but were completed for thoroughness. The average sequence coverage for the anadromous males was $0.36\times$ per individual and was $0.43\times$ per individual for the resident males (**Figure 3**). After filtering, a total of 5.64 million SNPs were included for further tests for differences in allele frequency between phenotypic groups. Analyses with the PoolParty pipeline included results for F_{ST} , FET, sliding window F_{ST} , and local score analyses. Analyses of allele frequencies among the resident and anadromous groups indicated that six markers on multiple chromosomes were significant after Bonferroni correction with FET (**Figure 4A**) including those on chromosomes Omy06, Omy08, Omy12, Omy26, and Omy29 (**Table 1**). However,



only SNPs on chromosome Omy12 were significant among groups with sliding window F_{ST} and were considered the strongest candidate region for anadromy (**Figure 4B**). Significant SNPs on chromosome Omy12 were located 28.6–43.6 kb from the nearest gene, NACHT, LRR, and PYD domain-containing protein 3 (NLRP3), detected with the sliding window F_{ST} analysis (**Table 1**). The gene positioned closest (10.9 kb) to the significant SNP detected with FET analysis on chromosome Omy12 was the COP9 signalosome complex subunit 6 (CSN6) (**Table 1**). Analyses of combined sexes for resident vs. anadromous (with highly imbalanced sex ratios) yielded a strong signal near the sdY sex determining region on chromosome Omy29 as expected (**Supplementary Figure S5**).

Further analyses that account for physically linked SNPs (local score) or polygenic association (RF) provided no consistent signals of association and thus were considered false positive or unreliable results. Analyses with local score revealed many small but significant peaks on every chromosome of the genome

(**Supplementary Figure S1**). Because local score is highly susceptible to false positives due to multiple testing effects of linked SNPs, a stringent critical value (p -value < 0.001 after Bonferroni correction) was applied and indicated that significant peaks occurred on several chromosomes (**Supplementary Figure S2**). We also searched for polygenic association for epistatic loci with RF methods and significant loci were detected in multiple trees. However, candidates were not repeatable across replicate trees in the analyses and thus were discounted as reliable signals of polygenic association (**Supplementary Figure S4**). The QQ plot displayed most SNPs falling along the line representative of the null hypothesis or SNPs with no association to the anadromy trait (**Supplementary Figure S6**). The QQ plot also had loci that display deviation from the null hypothesis line, including every chromosome Omy12 marker identified with the other analyses, which were clustered even further from the null hypothesis line and above the stringent significance threshold of observed $sF_{ST} > 0.5$ (**Supplementary Figure S6**).

TABLE 1 | Genomic position and closest gene annotations for significant markers from Fisher's Exact Test (FET) and sliding window Fst (sFst) analyses on chromosome (Chr) Omy12.

SNP	Chr	Position	FET p-value value	sF _{ST} Differentiation value	Source	Closest Gene(s)	Possible Gene Function
1326266	6	95913987	4.03E-10	NA	PoolParty - fet	Leucine-rich repeat extensin-like protein 1	Leucine-rich repeat
1661566	8	43382768	1.14E-09	NA	PoolParty - fet	MYC proto-oncogene, bHLH transcription factor b	Leucine zipper, Myc, Myc-type, basic helix-loop-helix (bHLH) domain, transcription regulator Myc, N-terminal
2557802	12	58504189	1.10E-09	NA	PoolParty - fet	COP9 signalosome complex subunit 6	Posttranslational modification, protein turnover, chaperones, cell wall/membrane/envelope biogenesis
4778210	26	1643096	2.62E-13	NA	PoolParty - fet	Calcineurin-like phosphoesterase domain, ApaH type	Carbohydrate transport and metabolism, posttranslational modification, protein turnover, chaperones, mobilome, prophages, transposons
5119622	29	77226	8.99E-10	NA	PoolParty - fet	Anoctamin	Protein dimerization activity, integral component of membrane
5119623	29	77228	4.23E-09	NA	PoolParty - fet	Anoctamin	Protein dimerization activity, integral component of membrane
204161	12	93625000	NA	0.487	PoolParty - Sfst	NACHT, LRR, and PYD domain-containing protein 3	Leucine-rich repeat, transcription, signal transduction mechanisms, inorganic ion transport and metabolism
204162	12	93630000	NA	0.573	PoolParty - Sfst	NACHT, LRR, and PYD domain-containing protein 3	Leucine-rich repeat, transcription, signal transduction mechanisms, inorganic ion transport and metabolism
204163	12	93635000	NA	0.579	PoolParty - Sfst	NACHT, LRR, and PYD domain-containing protein 3	Leucine-rich repeat, transcription, signal transduction mechanisms, inorganic ion transport and metabolism
204164	12	93640000	NA	0.579	PoolParty - Sfst	NACHT, LRR, and PYD domain-containing protein 3	Leucine-rich repeat, transcription, signal transduction mechanisms, inorganic ion transport and metabolism

TABLE 2 | Chromosome (Chr) Omy05 markers from Pearse et al. (2019) and candidate genetic markers from Narum et al. (2011) were analyzed with minor allele frequencies (MAF) and Fisher's Exact Test (FET) for the samples for this study.

Genetic marker name	Chr	Total MAF	Anadromous MAF	Resident MAF	Homozygous MAF	Heterozygous MAF	Homozygous Major Allele Frequency	Inversion	FET p-value
OmyR19198Pearse	5	0.021	0.026	0.019	0	0.043	0.957	1	1
OmyR40252Pearse	5	0.021	0.026	0.019	0	0.043	0.957	1	1
OmyR14589Pearse	5	0.021	0.026	0.019	0	0.043	0.957	1	1
Omy_RAD23894-58	5	0.137	0.179	0.115	0.009	0.256	0.735	1	0.202
OmyR24370Pearse	5	0.021	0.026	0.019	0	0.043	0.957	1	1
Omy_bcAKala-380th	5	0.462	0.449	0.468	0.248	0.427	0.325	2	0.799
OmyR33562Pearse	5	0.021	0.026	0.019	0	0.043	0.957	2	1
Omy_u09-61.043	5	0.000	0.000	0.000	0.000	0.000	1.000	2	No score, fixed
Omy_RAD30392-17	5	0.402	0.385	0.410	0.214	0.376	0.393	2	0.699
Omy_SECC22b-88	5	0.009	0.000	0.013	0	0.017	0.983	2	0.551
OmyR40319Pearse	5	0.004	0.000	0.006	0	0.009	0.991	2	1
Omy_ndk152	12	0.177	0.244	0.154	0.017	0.333	0.632	NA	0.13444
Omy_IL6320	14	0.177	0.205	0.173	0.060	0.248	0.692	NA	0.62277
Omy_LDHB2_i6	21	0.115	0.115	0.115	0.017	0.205	0.761	NA	0.83429

Results for the 12 markers in the inversion complex of chromosome Omy05 (Pearse et al., 2019) revealed that several were nearly fixed with very low variation in our samples, and others had no significant differences in allele frequencies (Table 2). The Omy_u09-61.043 marker was fixed for the

resident or derived allele, but this was not the case for all markers in this inversion (Table 2). Of the three candidate markers from Narum et al. (2011), two had similar allele frequencies between resident and anadromous groups, and none were significant after a FDR correction (Benjamini and Hochberg

1995). However, the marker with the lowest p -value ($p = 0.134$; Omy_ndk-152; Table 2) had an allele frequency difference of 9% between resident and anadromous groups and is located at position 59,538,315 on chromosome Omy12. This is near the region on chromosome Omy12 that was found to be significant in the whole-genome analyses.

4 DISCUSSION

In this study, we conducted multiple analyses to test for genomic regions of high differentiation between the resident and anadromous samples and examine consistent signals that would identify candidate genes associated with this life history trait. Tracking migratory behavior of PIT-tagged juveniles provided resident vs. anadromous (outmigration) phenotypes for genomic analyses, and distinct patterns in migration behavior were observed by each sex. The natural origin *O. mykiss* captured in the Klickitat River screw trap in 2018, 2019, and 2021 were classified as smolts due to distinct physical characteristics at rates of 99.5% ($n = 3,615$), 99.4% ($n = 3,475$), and 99.7% ($n = 2,352$), respectively. Outmigrating individuals detected in the Columbia River after leaving the Klickitat River were most likely steelhead smolts, but downstream migrating resident rainbow trout cannot be completely ruled out. Females rarely remained as residents in contrast to males despite collection from the same stream reaches, which is consistent with the differences in migratory behavior among sexes that has been generally observed in previous studies of this species (Fleming and Reynolds 2004; Quinn et al., 2011). The lack of female resident fish limited genomic analyses to largely focus on variation in male migratory phenotypes, but results revealed a candidate region of minor effect on chromosome Omy12.

Analyses with several approaches were considered necessary to account for limitations of each method and increase confidence in potential candidate genes. For example, F_{ST} and sliding window F_{ST} have been demonstrated to be susceptible to false positives or skewed signal to noise (Fariello et al., 2017), and thus, local score was applied from FET results to improve signal to noise in this high-density genome scan. The RF method was expected to be more effective to detect polygenetic signals of minor effect genes across the genome, especially for natural populations with low linkage disequilibrium (LD) (Holliday et al., 2012). However, analyses from this study yielded highest confidence from tests of sliding window F_{ST} and significance based on FET, whereas other results appeared to suffer from high levels of false positives.

Analyses provided highest confidence for a genomic region of minor effect on chromosome Omy12 for anadromy in male steelhead in the Klickitat River. We chose to align resequencing data from this study to the highest quality genome assembly that was currently available for the species (Gao et al., 2021). Although karyotypes vary across lineages of *O. mykiss*, the number of chromosome arms is constant (Thorgaard et al., 1983; Phillips et al., 2006; Pearse et al., 2019). The newest Arlee assembly has 32 haploid chromosomes due to fissions at chromosomes Omy04, Omy14, and Omy25, but chromosome Omy12 demonstrates homology between Arlee and Swanson lines (Gao et al., 2021).

Thus, we do not expect that results observed at Omy12 in the current study were due to problems in homology with the reference genome assembly. The gene nearest to significant SNPs was CSN6 has been shown to be overexpressed in a variety of cancer specimens (Zhang et al., 2013). Another nearby gene on chromosome Omy12 was the NLRP3, which has previously been observed functioning to restrict bacterial infection in Japanese flounder (Chen et al., 2020). The function of these genes did not bear an obvious biological connection to resident or anadromous behaviors and would require further validation before speculating on their related functions in *O. mykiss*.

Six markers specifically developed from the large inversion region on chromosome Omy05 (Pearse et al., 2019) had nearly zero variation among samples in this study and thus were not informative for analyses. This is consistent with results from Pearse et al. (2019) that observed significant results at chromosome Omy05 for coastal populations from the southern portion of the range but were not informative in other portions of the species' range. Additional markers that map to the same region on chromosome Omy05 that were developed in previous restriction site-associated DNA studies (Micheletti and Narum 2018) were variable but not significantly different between the two life history types in the Klickitat River. The difference in variation from markers in the chromosome Omy05 region was unexpected because there is extensive LD across the entire double inversion (Pearse et al., 2019), but this may reflect near fixation of derived alleles followed by subsequent recombination in this population that is part of the distinct inland lineage of *O. mykiss* (e.g., Collins et al., 2020). Previous candidate markers for anadromy in *O. mykiss* from the Klickitat River (Narum et al., 2011) were also not significant, but it was notable that one marker (Omy_ndk-154) was relatively near the same region on chromosome Omy12 identified in the whole-genome scan and had modest difference in allele frequencies between life history types. However, this marker may be too distant (~37 Mb) from the candidate region on chromosome Omy12 to be informative to distinguish life history variation. Thus, there was no evidence in this study population for association with Omy05 migratory behavior as has been shown in a previous study (Pearse et al., 2019). When the results of this study were compared to similar studies (Narum et al., 2011; Hecht et al., 2012; Hale et al., 2013; Hecht et al., 2013; Campbell et al., 2021), a significant association with anadromy was found on chromosome Omy12 in all cases, but these studies have consistently suggested polygenic association located throughout the genome.

Overall, analyses conducted on the *O. mykiss* samples collected from the Klickitat River for this study detected loci significantly associated with the trait of anadromy that may represent candidate genes of minor effect within this *O. mykiss* population. The significant region on chromosome Omy12 provided the strongest support; however, this region contains multiple candidate genes with uncertain biological relevance to resident versus anadromous life history types. Thus, candidate genes of major effect for anadromy remain elusive and results continue to support the hypothesis that the genetic basis for life history types in this lineage of *O. mykiss* is highly polygenic with minor effect (e.g., Nichols et al., 2008; Narum et al., 2011; Hecht

et al., 2012; Hale et al., 2013; Hecht et al., 2013; Campbell et al., 2021). Previous studies evaluated gene expression associated with anadromy at various stages of development of anadromous steelhead and resident rainbow trout and uncovered differentially expressed genes throughout the genome, including many on chromosome Omy12 (McKinney et al., 2015; Hale et al., 2016). In addition, Baerwald et al. (2016) evaluated differentially methylated regions (DMRs) to assess the role of epigenetics on smoltification in *O. mykiss* and identified one DMR associated with smoltification on chromosome Omy12. It is also likely that the main drivers of anadromy in this lineage of *O. mykiss* may be more strongly related to environmental rearing such as epigenetic modification, gene expression, and phenotypic plasticity. Future studies into these various mechanisms regulating anadromy are needed to further understand variation in the life history of this protected species.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: PRJNA771561.

ETHICS STATEMENT

The animal study was reviewed and approved by Monitoring; Methods (monitoringresources.org).

REFERENCES

- Baerwald, M. R., Meek, M. H., Stephens, M. R., Nagarajan, R. P., Goodbla, A. M., Tomalty, K. M. H., et al. (2016). Migration-related Phenotypic Divergence Is Associated with Epigenetic Modifications in Rainbow trout. *Mol. Ecol.* 25 (8), 1785–1800. doi:10.1111/mec.13231
- Benestan, L., Moore, J.-S., Sutherland, B. J. G., Le Luyer, J., Maaroufi, H., Rougeux, C., et al. (2017). Sex Matters in Massive Parallel Sequencing: Evidence for Biases in Genetic Parameter Estimation and Investigation of Sex Determination Systems. *Mol. Ecol.* 26 (24), 6767–6783. doi:10.1111/mec.14217
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B (Methodological)* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Blankenship, S. M., Campbell, M. R., Hess, J. E., Hess, M. A., Kassler, T. W., Kozfkay, C. C., et al. (2011). Major Lineages and Metapopulations in Columbia River *Oncorhynchus mykiss* Are Structured by Dynamic Landscape Features and Environments. *Trans. Am. Fish. Soc.* 140 (3), 665–684. doi:10.1080/00028487.2011.584487
- Brieuc, M. S. O., Waters, C. D., Drinan, D. P., and Naish, K. A. (2018). A Practical Introduction to Random Forest for Genetic Association Studies in Ecology and Evolution. *Mol. Ecol. Resour.* 18 (4), 755–766. doi:10.1111/1755-0998.12773
- Brunelli, J. P., Wertzler, K. J., Sundin, K., and Thorgaard, G. H. (2008). Y-specific Sequences and Polymorphisms in Rainbow trout and Chinook salmon. *Genome* 51 (9), 739–748. doi:10.1139/G08-060
- Bushnell, B. (2016). *BMap Short Read Aligner*. Berkeley, California: University of California <http://sourceforge.net/projects/bmap>.
- Campbell, M. A., Anderson, E. C., Garza, J. C., and Pearse, D. E. (2021). Polygenic Basis and the Role of Genome Duplication in Adaptation to Similar Selective Environments. *J. Hered.* 112 (7), 614–625. doi:10.1093/jhered/esab049
- Campbell, N. R., Harmon, S. A., and Narum, S. R. (2015). Genotyping-in-Thousands by Sequencing (GT-seq): A Cost Effective SNP Genotyping

AUTHOR CONTRIBUTIONS

SN and JZ directed the study. SN, JZ, and NR designed the study. EC analyzed the data. All authors interpreted the results and wrote the manuscript.

FUNDING

Funding for this project was provided by the Bonneville Power Administration through grant number 199506335 to Yakima Nation.

ACKNOWLEDGMENTS

We thank YKFP for providing samples (particularly, Kory Kuhn, Sandy Pinkham, Rod Begay, and Dean Antone) and the CRITFC laboratory staff involved in sample processing (particularly, Amber Gonzalez, Megan Moore, Lori Maxwell, Stephanie Harmon, Joe Gasper, and Kim Vertacnik).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.795850/full#supplementary-material>

- Method Based on Custom Amplicon Sequencing. *Mol. Ecol. Resour.* 15 (4), 855–867. doi:10.1111/1755-0998.12357
- Chen, H., Ding, S., Tan, J., Yang, D., Zhang, Y., and Liu, Q. (2020). Characterization of the Japanese Flounder NLRP3 Inflammasome in Restricting *Edwardsiella piscicida* Colonization *In Vivo*. *Fish Shellfish Immunol.* 103, 169–180. doi:10.1016/j.fsi.2020.04.063
- Christie, M. R., Marine, M. L., and Blouin, M. S. (2011). Who Are the Missing Parents? Grandparentage Analysis Identifies Multiple Sources of Gene Flow into a Wild Population. *Mol. Ecol.* 20, 1263–1276. doi:10.1111/j.1365-294X.2010.04994.x
- Collins, E. E., Hargrove, J. S., Delomas, T. A., and Narum, S. R. (2020). Distribution of Genetic Variation Underlying Adult Migration Timing in Steelhead of the Columbia River basin. *Ecol. Evol.* 1017, 9486–9502. doi:10.1002/ece3.6641
- Courter, I. I., Child, D. B., Hobbs, J. A., Garrison, T. M., Glessner, J. J. G., and Duery, S. (2013). Resident Rainbow trout Produce Anadromous Offspring in a Large interior Watershed. *Can. J. Fish. Aquat. Sci.* 70 (5), 701–710. doi:10.1139/cjfas-2012-0457
- Crawford, B. (1979). *The Origin and History of the trout Brood Stocks of the Washington Department of Game*. Olympia: Washington State Game Department, Fisheries Research Report.
- Evans, S. D., Lindley, D. S., Kock, T. J., Hansen, A. C., Perry, R. W., Zendt, J. S., et al. (2021). *Evaluation of Movement and Survival of Juvenile Steelhead (Oncorhynchus mykiss) and Coho salmon (Oncorhynchus kisutch) in the Klickitat River, Washington, 2018–2019 (No. 2021-1083)*. Reston, Virginia, U.S.: US Geological Survey.
- Fariello, M. I., Boitard, S., Mercier, S., Robelin, D., Faraut, T., Arnould, C., et al. (2017). Accounting for Linkage Disequilibrium in Genome Scans for Selection without Individual Genotypes: the Local Score Approach. *Mol. Ecol.* 26 (14), 3700–3714. doi:10.1111/mec.14141
- Faust, G. G., and Hall, I. M. (2014). SAMBLASTER: Fast Duplicate Marking and Structural Variant Read Extraction. *Bioinformatics* 30 (17), 2503–2505. doi:10.1093/bioinformatics/btu314 <https://github.com/GregoryFaust/samblaster>

- Fleming, I. A., and Reynolds, J. D. (2004). "Salmonid Breeding Systems," in *Evolution Illuminated: salmon and Their Relatives*. Editors A. P. Hendry and S. C. Stearns (Oxford, UK: Oxford University Press), 264–294.
- Gao, G., Magadan, S., Waldbieser, G. C., Youngblood, R. C., Wheeler, P. A., Scheffler, B. E., et al. (2021). A Long Reads-Based De-novo Assembly of the Genome of the Arlee Homozygous Line Reveals Chromosomal Rearrangements in Rainbow trout. *G3-Genes Genom Genet* 11, 4. doi:10.1093/g3journal/jkab052
- Goldstein, B. A., Polley, E. C., and Briggs, F. B. S. (2011). Random Forests for Genetic Association Studies. *Stat. Appl. Genet. Mol.* 10, 35. doi:10.2202/1544-6115.1691
- Hale, M. C., McKinney, G. J., Thrower, F. P., and Nichols, K. M. (2016). RNA-seq Reveals Differential Gene Expression in the Brains of Juvenile Resident and Migratory Smolt Rainbow trout (*Oncorhynchus mykiss*). *Comp. Biochem. Physiol. D: Genomics Proteomics* 20, 136–150. doi:10.1016/j.cbcd.2016.07.006
- Hale, M. C., Thrower, F. P., Berntson, E. A., Miller, M. R., and Nichols, K. M. (2013). Evaluating Adaptive Divergence between Migratory and Nonmigratory Ecotypes of a Salmonid Fish, *Oncorhynchus mykiss*. *G3-genes Genom. Genet.* 3 (8), 1273–1285. doi:10.1534/g3.113.006817
- Hecht, B. C., Campbell, N. R., Holecsek, D. E., and Narum, S. R. (2013). Genome-wide Association Reveals Genetic Basis for the Propensity to Migrate in Wild Populations of Rainbow and Steelhead trout. *Mol. Ecol.* 22 (11), 3061–3076. doi:10.1111/mec.12082
- Hecht, B. C., Thrower, F. P., Hale, M. C., Miller, M. R., and Nichols, K. M. (2012). Genetic Architecture of Migration-Related Traits in Rainbow and Steelhead Trout, *Oncorhynchus mykiss*. *G3-genes Genom. Genet.* 29, 1113–1127. doi:10.1534/g3.112.003137
- Holliday, J. A., Wang, T., and Aitken, S. (2012). Predicting Adaptive Phenotypes from Multilocus Genotypes in Sitka spruce (*Picea Sitchensis*) Using Random forest. *G* 2, 1085–1093. doi:10.1534/g3.112.002733
- Horn, R. L., Kamphaus, C., Murdoch, K., and Narum, S. R. (2020). Detecting Genomic Variation Underlying Phenotypic Characteristics of Reintroduced Coho salmon (*Oncorhynchus kisutch*). *Conserv. Genet.* 21 (6), 1011–1021. doi:10.1007/s10592-020-01307-0
- Kofler, R., Pandey, R. V., and Schlotterer, C. (2011). PoPoolation2: Identifying Differentiation between Populations Using Sequencing of Pooled DNA Samples (Pool-Seq). *Bioinformatics* 27 (24), 3435–3436. doi:10.1093/bioinformatics/btr589
- Li, H. (2013). Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM. arXiv [Preprint]. Available at: <https://arxiv.org/pdf/1303.3997.pdf> (Accessed September 27, 2021 <http://github.com/lh3/bwa>).
- Li, J., Phillips, R. B., Harwood, A. S., Koop, B. F., and Davidson, W. S. (2011). Identification of the Sex Chromosomes of Brown trout (*Salmo trutta*) and Their Comparison with the Corresponding Chromosomes in Atlantic salmon (*Salmo salar*) and Rainbow trout (*Oncorhynchus mykiss*). *Cytogenet. Genome Res.* 133 (1), 25–33. doi:10.1159/000323410
- Liauw, A., and Wiener, M. (2002). Classification and Regression by randomForest. *R. News* 2 (3), 18–22. <http://www.stat.Berkeley.edu/users/breiman/>
- Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: Genome Association and Prediction Integrated Tool. *Bioinformatics* 28 (18), 2397–2399. doi:10.1093/bioinformatics/bts444 <http://www.maizegenetics.net/GAPI>
- Martínez, A., Garza, J. C., and Pearse, D. E. (2011). A Microsatellite Genome Screen Identifies Chromosomal Regions under Differential Selection in Steelhead and Rainbow trout. *Trans. Am. Fish. Soc.* 140 (3), 829–842. doi:10.1080/00028487.2011.588094
- McDowall, R. M. (1988). *Diadromy in Fishes*. Portland, OR, US: Timber Press.
- McKinney, G. J., Hale, M. C., Goetz, G., Gribskov, M., Thrower, F. P., and Nichols, K. M. (2015). Ontogenetic Changes in Embryonic and Brain Gene Expression in Progeny Produced from Migratory and resident *Oncorhynchus Mykiss*. *Mol. Ecol.* 24 (8), 1792–1809. doi:10.1111/mec.13143
- Micheletti, S. J., and Narum, S. R. (2018). Utility of Pooled Sequencing for Association Mapping in Nonmodel Organisms. *Mol. Ecol. Resour.* 18 (4), 825–837. doi:10.1111/1755-0998.1278 <https://github.com/StevenMicheletti/poolparty>.
- Moore, J. W., Yeakel, J. D., Peard, D., Lough, J., and Beere, M. (2014). Life-history Diversity and its Importance to Population Stability and Persistence of a Migratory Fish: Steelhead in Two Large North American Watersheds. *J. Anim. Ecol.* 83 (5), 1035–1046. doi:10.1111/1365-2656.12212
- Narum, S. R., Powell, M. S., Evenson, R., Sharp, B., and Talbot, A. (2006). Microsatellites Reveal Population Substructure of Klickitat River Native Steelhead and Genetic Divergence from an Introduced Stock. *North Am. J. Fish. Manage.* 26 (1), 147–155. doi:10.1577/M05-055.1
- Narum, S. R., Zendt, J. S., Frederiksen, C., Campbell, N., Matala, A., and Sharp, W. R. (2011). Candidate Genetic Markers Associated with Anadromy In *Oncorhynchus mykiss* of the Klickitat River. *Trans. Am. Fish. Soc.* 140 (3), 843–854. doi:10.1080/00028487.2011.588131
- Narum, S. R., Zendt, J. S., Graves, D., and Sharp, W. R. (2008). Influence of Landscape on Resident and Anadromous Life History Types of *Oncorhynchus mykiss*. *Can. J. Fish. Aquat. Sci.* 65 (6), 1013–1023. doi:10.1139/F08-025
- Nichols, K. M., Edo, A. F., Wheeler, P. A., and Thorgaard, G. H. (2008). The Genetic Basis of Smoltification-Related Traits in *Oncorhynchus mykiss*. *Genetics* 179 (3), 1559–1575. doi:10.1534/genetics.107.084251
- Pearse, D. E., Barson, N. J., Nome, T., Gao, G., Campbell, M. A., Abadía-Cardoso, A., et al. (2019). Sex-dependent Dominance Maintains Migration Supergene in Rainbow trout. *Nat. Ecol. Evol.* 3 (12), 1731–1742. doi:10.1038/s41559-019-1044-6
- Phillips, R. B., Nichols, K. M., DeKoning, J. J., Morasch, M. R., Keatley, K. A., Rexroad III, C., et al. (2006). Assignment of Rainbow Trout Linkage Groups to Specific Chromosomes. *Genetics* 174 (3), 1661–1670. doi:10.1534/genetics.105.055269
- Quinn, T. P., Seamons, T. R., Vøllestad, L. A., and Duffy, E. (2011). Effects of Growth and Reproductive History on the Egg Size-Fecundity Tradeoff in Steelhead. *Trans. Am. Fish. Soc.* 140, 45–51. doi:10.1080/00028487.2010.550244
- Rokach, L. (2016). Decision forest: Twenty Years of Research. *Inf. Fusion* 27, 111–125. doi:10.1016/j.inffus.2015.06.005
- Sloat, M. R., and Reeves, G. H. (2014). Individual Condition, Standard Metabolic Rate, and Rearing Temperature Influence Steelhead and Rainbow trout (*Oncorhynchus mykiss*) Life Histories. *Can. J. Fish. Aquat. Sci.* 71 (4), 491–501. doi:10.1139/cjfas-2013-0366
- Sweet, D., Lorente, M., Valenzuela, A., Lorente, J., and Alvarez, J. C. (1996). Increasing DNA Extraction Yield from Saliva Stains with a Modified Chelex Method. *Forensic Sci. Int.* 83 (3), 167–177. doi:10.1016/s0379-0738(96)02034-8
- Thorgaard, G. H. (1983). Chromosomal Differences Among Rainbow Trout Populations. *Copeia*, 650–662. doi:10.2307/1444329
- VanRaden, P. M., Olson, K. M., Wiggans, G. R., Cole, J. B., and Tooker, M. E. (2011). Genomic Inbreeding and Relationships Among Holsteins, Jerseys, and Brown Swiss. *J. Dairy Sci.* 94, 5673–5682. doi:10.3168/jds.2011-4500
- Waples, R. S., Zabel, R. W., Scheuerell, M. D., and Sanderson, B. L. (2008). Evolutionary Responses by Native Species to Major Anthropogenic Changes to Their Ecosystems: Pacific salmon in the Columbia River Hydropower System. *Mol. Ecol.* 17, 84–96. doi:10.1111/j.1365-294X.2007.03510.x
- Yano, A., Nicol, B., Jouanno, E., Quillet, E., Fostier, A., Guyomard, R., et al. (2013). The Sexually Dimorphic on the Y-chromosome Gene (sdY) Is a Conserved Male-specific Y-chromosome Sequence in many Salmonids. *Evol. Appl.* 6 (3), 486–496. doi:10.1111/eva.12032
- Zhang, S.-N., Pei, D.-S., and Zheng, J.-N. (2013). The COP9 Signalosome Subunit 6 (CSN6): a Potential Oncogene. *Cell div* 8 (1), 1–5. doi:10.1186/1747-1028-8-14
- Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed Linear Model Approach Adapted for Genome-wide Association Studies. *Nat. Genet.* 42, 355–360. doi:10.1038/ng.546

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Collins, Romero, Zendt and Narum. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Supplementary Table 1. Sample names, life history types, sample tributary, and years.

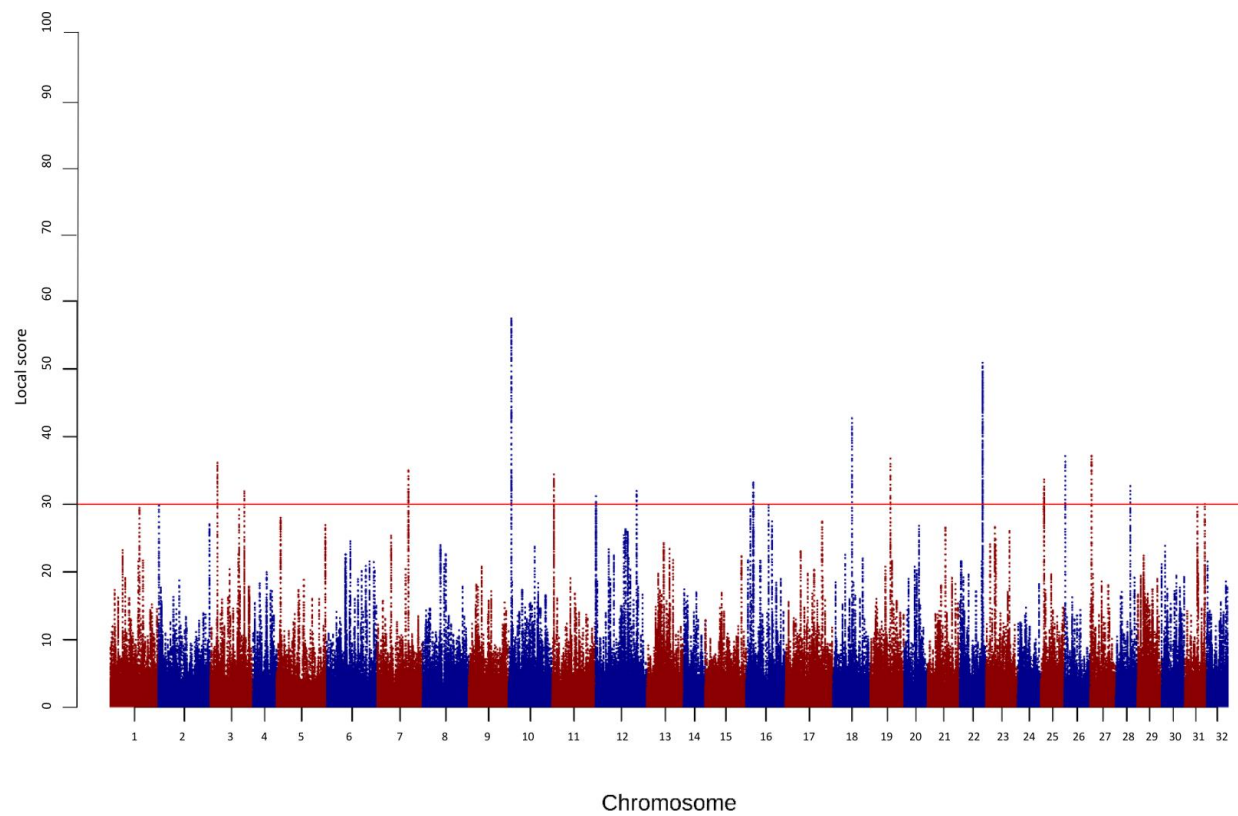
Sample Name	Life history type	Tributary	Latitude	Longitude	Year	Sex	Retained for final analyses?
OmyKR11-BRC08-C1	Anadromous	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-BRC08-C2	Anadromous	Brush Creek	46.163	-120.945	2011	F	N
OmyKR11-BRC08-C3	Resident	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-BRC08-E8	Resident	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-BRC08-F2	Anadromous	Brush Creek	46.163	-120.945	2011	F	N
OmyKR11-BRC08-G2	Resident	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-BRC08-I5	Resident	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-BRC38-A3	Resident	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-BRC41-H5	Resident	Brush Creek	46.163	-120.945	2011	M	Y
OmyKR11-ETC13-B3	Resident	EF Tepee Creek	46.155	-121.020	2011	M	Y
OmyKR11-TPC10-B5	Anadromous	Tepee Creek	46.181	-121.025	2011	F	N
OmyKR11-TPC10-C9	Anadromous	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC10-E1	Anadromous	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC10-E3	Anadromous	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC11-C1	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC11-G9	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC12-B6	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC17-A1	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC17-C6	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC17-E4	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC20-C2	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC21-A10	Anadromous	Tepee Creek	46.181	-121.025	2011	F	N
OmyKR11-TPC21-A4	Anadromous	Tepee Creek	46.181	-121.025	2011	F	N
OmyKR11-TPC22-C4	Anadromous	Tepee Creek	46.181	-121.025	2011	F	N
OmyKR11-TPC23-A2	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC24-A6	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC24-B1	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC24-B4	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC24-B5	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC24-D3	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC37-A8	Anadromous	Tepee Creek	46.181	-121.025	2011	F	N
OmyKR11-TPC37-B6	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC48-B7	Resident	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-TPC50-B10	Anadromous	Tepee Creek	46.181	-121.025	2011	M	Y
OmyKR11-WHC09-A10	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC15-G5	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC18-A5	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC19-B1	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC19-D3	Anadromous	White Creek	46.013	-121.150	2011	M	Y

OmyKR11-WHC19-F1	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC19-F10	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC19-G3	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC19-H3	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC25-B5	Anadromous	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC27-A4	Resident	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC27-B2	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC27-B3	Resident	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC28-B7	Resident	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC29-F6	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC30-I3	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC30-J6	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC31-C9	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC31-J10	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC32-A5	Anadromous	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC32-B3	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC32-D3	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC33-H9	Anadromous	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC34-I6	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC34-J4	Anadromous	White Creek	46.013	-121.150	2011	F	N
OmyKR11-WHC35-A4	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WHC40-B3	Resident	White Creek	46.013	-121.150	2011	M	Y
OmyKR11-WWC14-C2	Resident	WF White Creek	46.180	-121.070	2011	M	Y
OmyKR12-BLC12-A8	Anadromous	Blue Creek	46.076	-121.055	2012	F	N
OmyKR12-BRC02-D5	Resident	Brush Creek	46.163	-120.945	2012	M	Y
OmyKR12-BRC02-G1	Resident	Brush Creek	46.163	-120.945	2012	M	Y
OmyKR12-BRC02-G2	Resident	Brush Creek	46.163	-120.945	2012	M	Y
OmyKR12-BRC02-H7	Resident	Brush Creek	46.163	-120.945	2012	M	Y
OmyKR12-BRC03-E7	Resident	Brush Creek	46.163	-120.945	2012	M	Y
OmyKR12-BRC26-C9	Anadromous	Brush Creek	46.163	-120.945	2012	F	N
OmyKR12-TPC15-E7	Anadromous	Tepee Creek	46.181	-121.025	2012	F	N
OmyKR12-TPC15-J3	Resident	Tepee Creek	46.181	-121.025	2012	M	Y
OmyKR12-TPC40-A10	Anadromous	Tepee Creek	46.181	-121.025	2012	M	Y
OmyKR12-WHC13-G5	Resident	White Creek	46.013	-121.150	2012	M	Y
OmyKR12-WHC22-B10	Anadromous	White Creek	46.013	-121.150	2012	F	N
OmyKR12-WHC24-J8	Resident	White Creek	46.013	-121.150	2012	M	Y
OmyKR12-WHC31-G7	Anadromous	White Creek	46.013	-121.150	2012	F	N
OmyKR12-WHC41-A6	Anadromous	White Creek	46.013	-121.150	2012	F	N
OmyKR13-BLC29-D10	Resident	Blue Creek	46.076	-121.055	2013	M	Y
OmyKR13-BLC29-G7	Anadromous	Blue Creek	46.076	-121.055	2013	F	N
OmyKR13-BRC32-C6	Anadromous	Brush Creek	46.163	-120.945	2013	F	N
OmyKR13-BRC33-B1	Anadromous	Brush Creek	46.163	-120.945	2013	U	N
OmyKR13-TPC05-A5	Anadromous	Tepee Creek	46.181	-121.025	2013	F	N

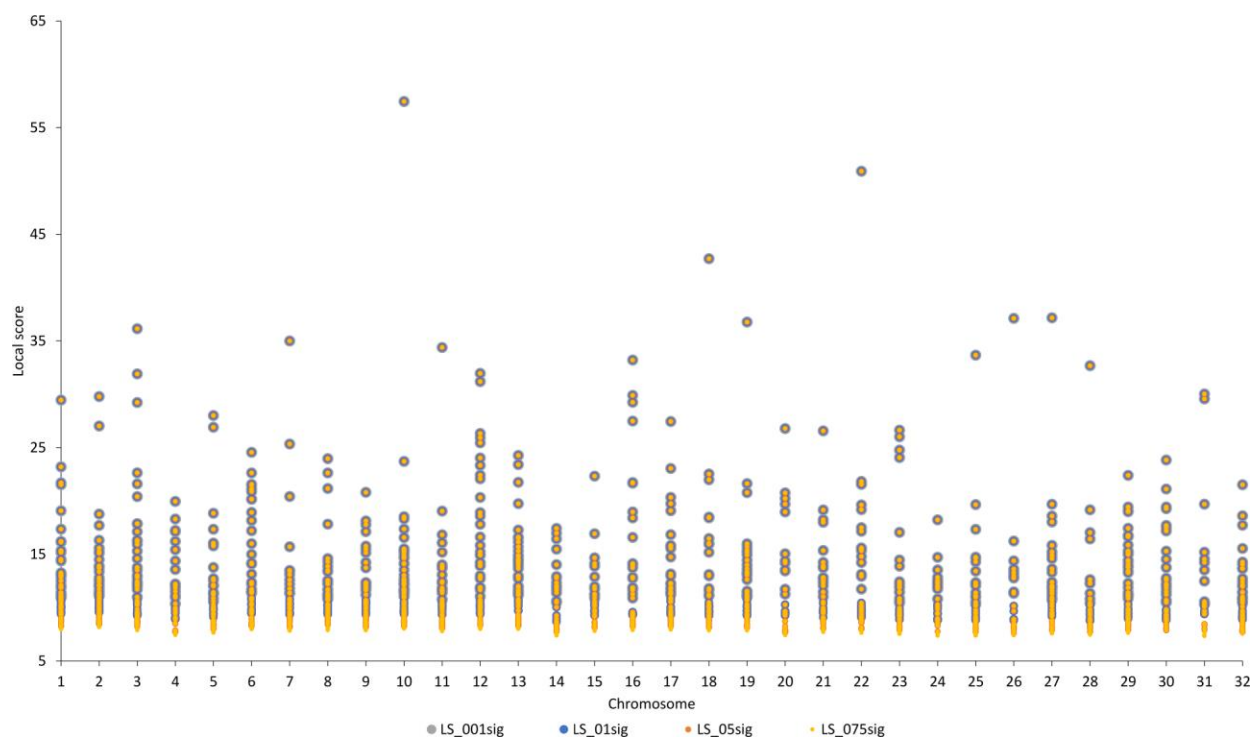
OmyKR13-TPC08-D9	Anadromous	Tepee Creek	46.181	-121.025	2013	M	Y
OmyKR13-TPC10-F6	Resident	Tepee Creek	46.181	-121.025	2013	M	Y
OmyKR13-TPC26-A2	Resident	Tepee Creek	46.181	-121.025	2013	M	Y
OmyKR13-TPC27-A1	Resident	Tepee Creek	46.181	-121.025	2013	M	Y
OmyKR13-WHC13-G5	Anadromous	White Creek	46.013	-121.150	2013	M	Y
OmyKR13-WHC14-D7	Resident	White Creek	46.013	-121.150	2013	M	Y
OmyKR13-WHC17-A9	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR13-WHC19-C2	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR13-WHC24-J5	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR13-WHC35-E5	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR13-WHC35-G4	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR13-WHC35-I3	Anadromous	White Creek	46.013	-121.150	2013	M	Y
OmyKR13-WHC37-A2	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR13-WHC37-H1	Anadromous	White Creek	46.013	-121.150	2013	F	N
OmyKR14-BRC12-F7	Anadromous	Brush Creek	46.163	-120.945	2014	F	N
OmyKR14-BRC12-G8	Anadromous	Brush Creek	46.163	-120.945	2014	F	N
OmyKR14-BRC33-C3	Anadromous	Brush Creek	46.163	-120.945	2014	M	Y
OmyKR14-ETC20-C6	Anadromous	EF Tepee Creek	46.155	-121.020	2014	F	N
OmyKR14-ETC20-C7	Anadromous	EF Tepee Creek	46.155	-121.020	2014	M	Y
OmyKR14-TPC14-A8	Resident	Tepee Creek	46.181	-121.025	2014	M	Y
OmyKR14-TPC14-G6	Anadromous	Tepee Creek	46.181	-121.025	2014	F	N
OmyKR14-TPC15-A3	Resident	Tepee Creek	46.181	-121.025	2014	M	Y
OmyKR14-TPC15-A9	Resident	Tepee Creek	46.181	-121.025	2014	M	Y
OmyKR14-TPC24-A2	Anadromous	Tepee Creek	46.181	-121.025	2014	M	Y
OmyKR14-TPC25-A3	Anadromous	Tepee Creek	46.181	-121.025	2014	M	Y
OmyKR14-TPC25-D3	Anadromous	Tepee Creek	46.181	-121.025	2014	F	N
OmyKR14-WHC34-A3	Resident	White Creek	46.013	-121.150	2014	M	Y
OmyKR14-WHC34-B2	Anadromous	White Creek	46.013	-121.150	2014	F	N
OmyKR14-WHC34-D7	Anadromous	White Creek	46.013	-121.150	2014	F	N
OmyKR14-WHC35-G2	Resident	White Creek	46.013	-121.150	2014	M	Y
OmyKR14-WHC35-J2	Anadromous	White Creek	46.013	-121.150	2014	M	Y
OmyKR14-WHC35-J9	Anadromous	White Creek	46.013	-121.150	2014	M	Y
OmyKR14-WHC37-G8	Resident	White Creek	46.013	-121.150	2014	M	Y
OmyKR14-WHC38-C3	Anadromous	White Creek	46.013	-121.150	2014	F	N
OmyKR14-WHC39-G3	Anadromous	White Creek	46.013	-121.150	2014	M	Y
OmyKR15-BRC20-A2	Anadromous	Brush Creek	46.163	-120.945	2015	F	N
OmyKR15-TPC21-C10	Anadromous	Tepee Creek	46.181	-121.025	2015	M	Y
OmyKR15-TPC25-G6	Anadromous	Tepee Creek	46.181	-121.025	2015	M	Y
OmyKR15-TPC25-G9	Anadromous	Tepee Creek	46.181	-121.025	2015	F	N
OmyKR15-TPC25-I1	Anadromous	Tepee Creek	46.181	-121.025	2015	F	N
OmyKR15-WHC20-C1	Resident	White Creek	46.013	-121.150	2015	M	Y
OmyKR15-WHC29-I7	Anadromous	White Creek	46.013	-121.150	2015	M	Y
OmyKR15-WHC31-A9	Resident	White Creek	46.013	-121.150	2015	M	Y

OmyKR16-BRC12-B7	Resident	Brush Creek	46.163	-120.945	2016	M	Y
OmyKR16-BRC12-E2	Anadromous	Brush Creek	46.163	-120.945	2016	F	N
OmyKR16-BRC20-C4	Anadromous	Brush Creek	46.163	-120.945	2016	F	N
OmyKR16-BRC28-D3	Resident	Brush Creek	46.163	-120.945	2016	M	Y
OmyKR16-TPC15-A2	Resident	Tepee Creek	46.181	-121.025	2016	M	Y
OmyKR16-TPC15-A4	Resident	Tepee Creek	46.181	-121.025	2016	M	Y
OmyKR16-TPC36-A5	Anadromous	Tepee Creek	46.181	-121.025	2016	F	N
OmyKR16-WHC19-C4	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR16-WHC19-C5	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR16-WHC23-A7	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR16-WHC23-D7	Anadromous	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC23-D8	Anadromous	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC23-G2	Anadromous	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC23-H5	Resident	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC23-I4	Resident	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC23-I9	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR16-WHC25-A5	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR16-WHC25-C3	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR16-WHC25-C5	Anadromous	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC26-A6	Resident	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC27-A7	Resident	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC27-F3	Resident	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC27-G8	Resident	White Creek	46.013	-121.150	2016	M	Y
OmyKR16-WHC29-C1	Anadromous	White Creek	46.013	-121.150	2016	F	N
OmyKR17-BRC02-B2	Resident	Brush Creek	46.163	-120.945	2017	M	N
OmyKR17-BRC02-B7	Resident	Brush Creek	46.163	-120.945	2017	M	Y
OmyKR17-BRC02-D6	Resident	Brush Creek	46.163	-120.945	2017	M	Y
OmyKR17-BRC04-B8	Resident	Brush Creek	46.163	-120.945	2017	M	Y
OmyKR17-BRC04-D6	Resident	Brush Creek	46.163	-120.945	2017	M	Y
OmyKR17-BRC04-E5	Resident	Brush Creek	46.163	-120.945	2017	M	Y
OmyKR17-TPC03-F6	Resident	Tepee Creek	46.181	-121.025	2017	M	Y
OmyKR17-TPC07-A9	Resident	Tepee Creek	46.181	-121.025	2017	M	Y
OmyKR17-TPC07-F10	Resident	Tepee Creek	46.181	-121.025	2017	M	Y
OmyKR17-WHC06-C8	Anadromous	White Creek	46.013	-121.150	2017	F	N
OmyKR17-WHC07-H3	Resident	White Creek	46.013	-121.150	2017	M	Y
OmyKR17-WHC07-H5	Resident	White Creek	46.013	-121.150	2017	M	Y
OmyKR17-WHC08-A6	Resident	White Creek	46.013	-121.150	2017	M	Y
OmyKR17-WHC08-F4	Resident	White Creek	46.013	-121.150	2017	M	Y
OmyKR17-WHC10-A10	Resident	White Creek	46.013	-121.150	2017	M	Y
OmyKR17-WHC10-E6	Resident	White Creek	46.013	-121.150	2017	M	Y
OmyKR18-BRC01-C8	Anadromous	Brush Creek	46.163	-120.945	2018	M	Y
OmyKR18-BRC02-D9	Anadromous	Brush Creek	46.163	-120.945	2018	F	N
OmyKR18-BRC04-B5	Anadromous	Brush Creek	46.163	-120.945	2018	F	N

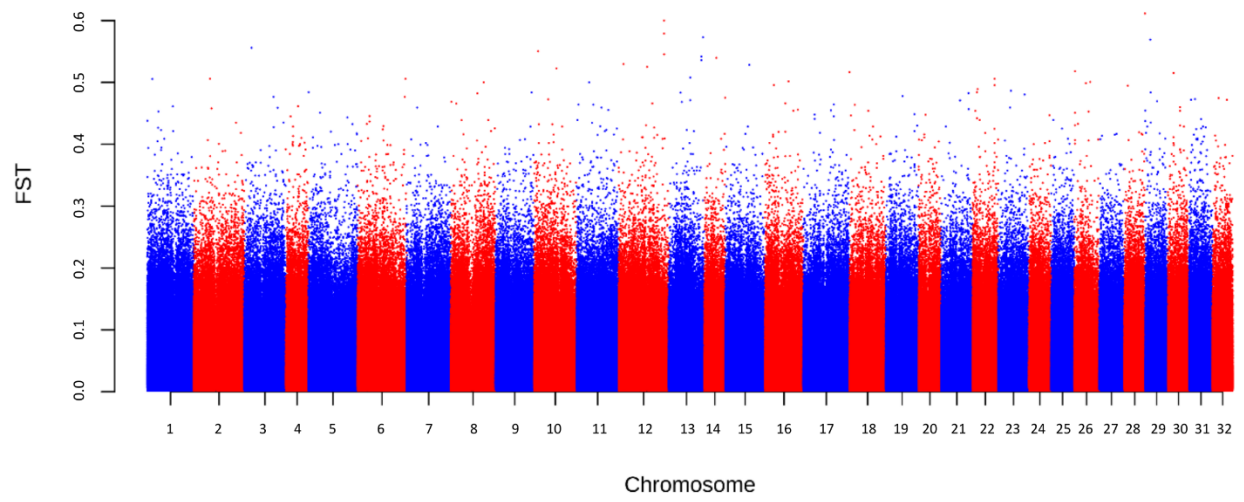
OmyKR18-BRC04-B6	Anadromous	Brush Creek	46.163	-120.945	2018	F	N
OmyKR18-BRC04-G5	Anadromous	Brush Creek	46.163	-120.945	2018	F	N
OmyKR18-BRC05-A4	Anadromous	Brush Creek	46.163	-120.945	2018	M	Y
OmyKR18-BRC05-B6	Anadromous	Brush Creek	46.163	-120.945	2018	F	N
OmyKR18-BRC05-D1	Anadromous	Brush Creek	46.163	-120.945	2018	F	N
OmyKR18-TPC01-A10	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC02-C5	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC02-D4	Anadromous	Tepee Creek	46.181	-121.025	2018	M	Y
OmyKR18-TPC02-E3	Anadromous	Tepee Creek	46.181	-121.025	2018	M	Y
OmyKR18-TPC02-G6	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC03-C2	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC03-F2	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC05-A9	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC05-B4	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC05-I9	Anadromous	Tepee Creek	46.181	-121.025	2018	M	Y
OmyKR18-TPC05-J6	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC06-A1	Anadromous	Tepee Creek	46.181	-121.025	2018	M	N
OmyKR18-TPC06-A8	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-TPC06-C8	Anadromous	Tepee Creek	46.181	-121.025	2018	M	Y
OmyKR18-TPC11-B4	Anadromous	Tepee Creek	46.181	-121.025	2018	F	N
OmyKR18-WHC04-E6	Anadromous	White Creek	46.013	-121.150	2018	F	N
OmyKR18-WHC06-H8	Anadromous	White Creek	46.013	-121.150	2018	F	N
OmyKR18-WHC08-B9	Anadromous	White Creek	46.013	-121.150	2018	M	Y
OmyKR18-WHC16-C9	Anadromous	White Creek	46.013	-121.150	2018	M	Y
OmyKR18-WHC19-D8	Anadromous	White Creek	46.013	-121.150	2018	F	N
OmyKR18-WHC19-F8	Anadromous	White Creek	46.013	-121.150	2018	M	Y
OmyKR18-WHC19-H9	Anadromous	White Creek	46.013	-121.150	2018	F	N
OmyKR18-WHC20-A7	Anadromous	White Creek	46.013	-121.150	2018	M	Y
OmyKR18-WHC20-G8	Anadromous	White Creek	46.013	-121.150	2018	M	Y
OmyKR19-WHC09-J6	Anadromous	White Creek	46.013	-121.150	2019	M	Y



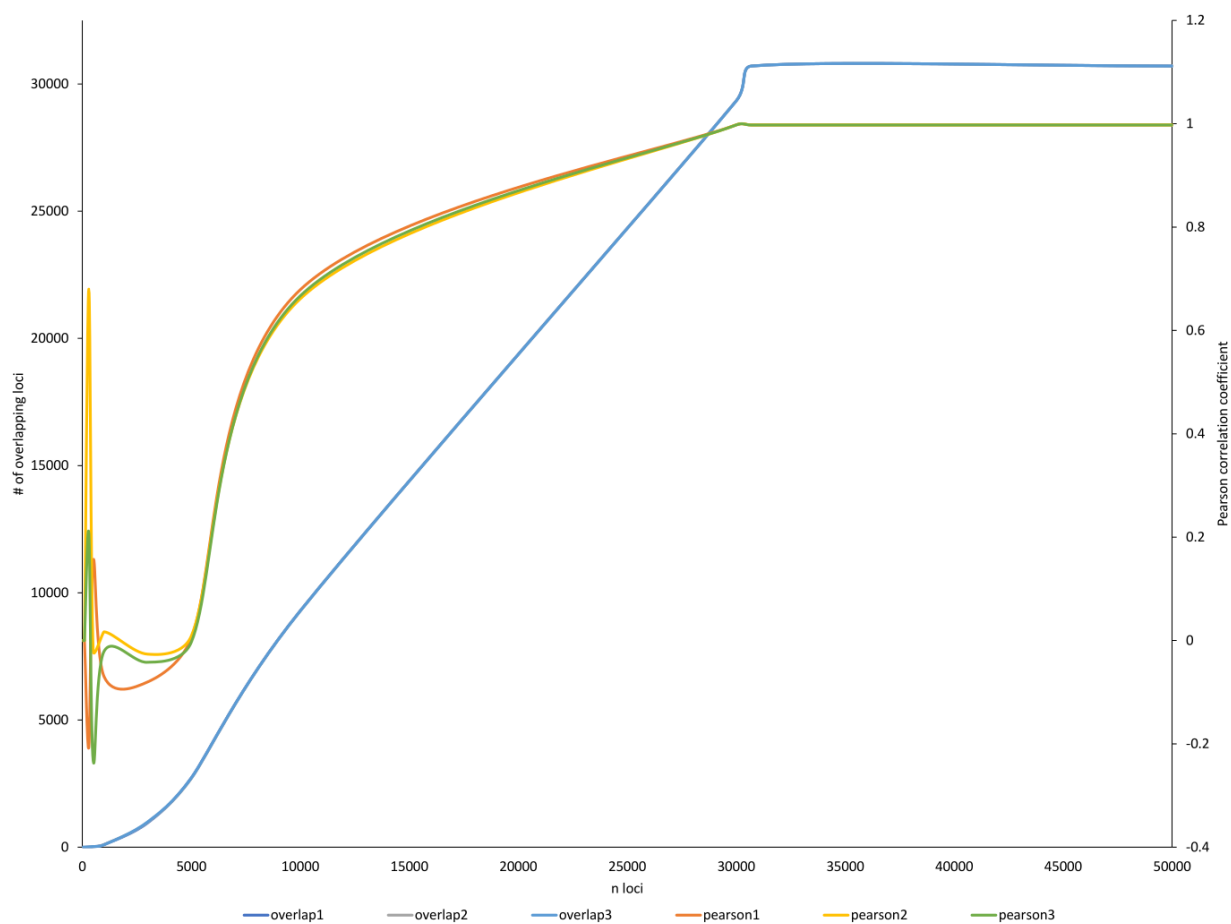
Supplementary Figure 1. Local score Manhattan plot. The minimum coverage threshold was 15 reads and the minor allele frequency was ≥ 0.05 .



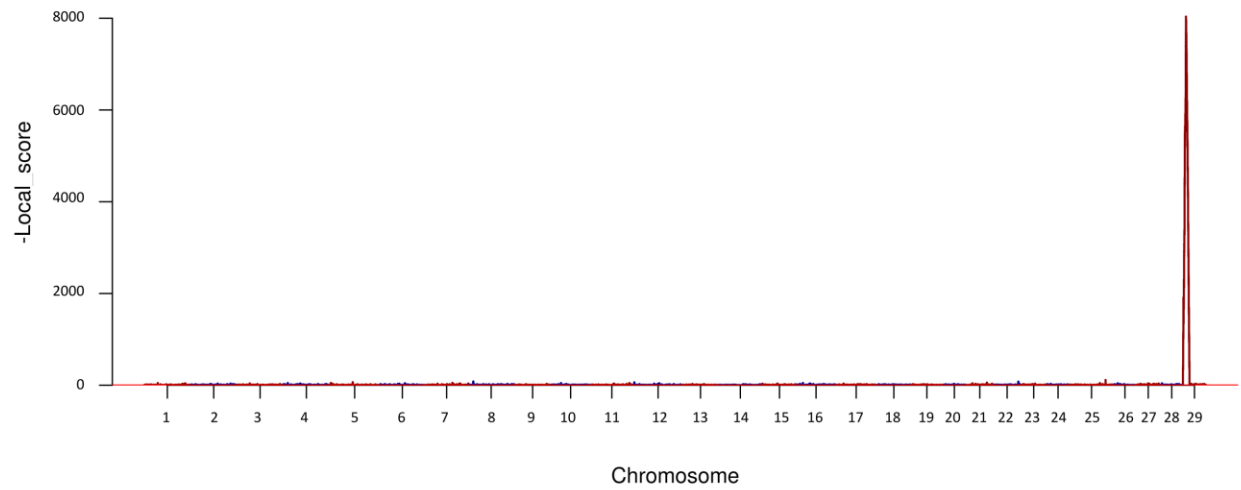
Supplementary Figure 2. Manhattan plot of significance of local scores. The minimum coverage threshold was 15 reads and the minor allele frequency was ≥ 0.05 .



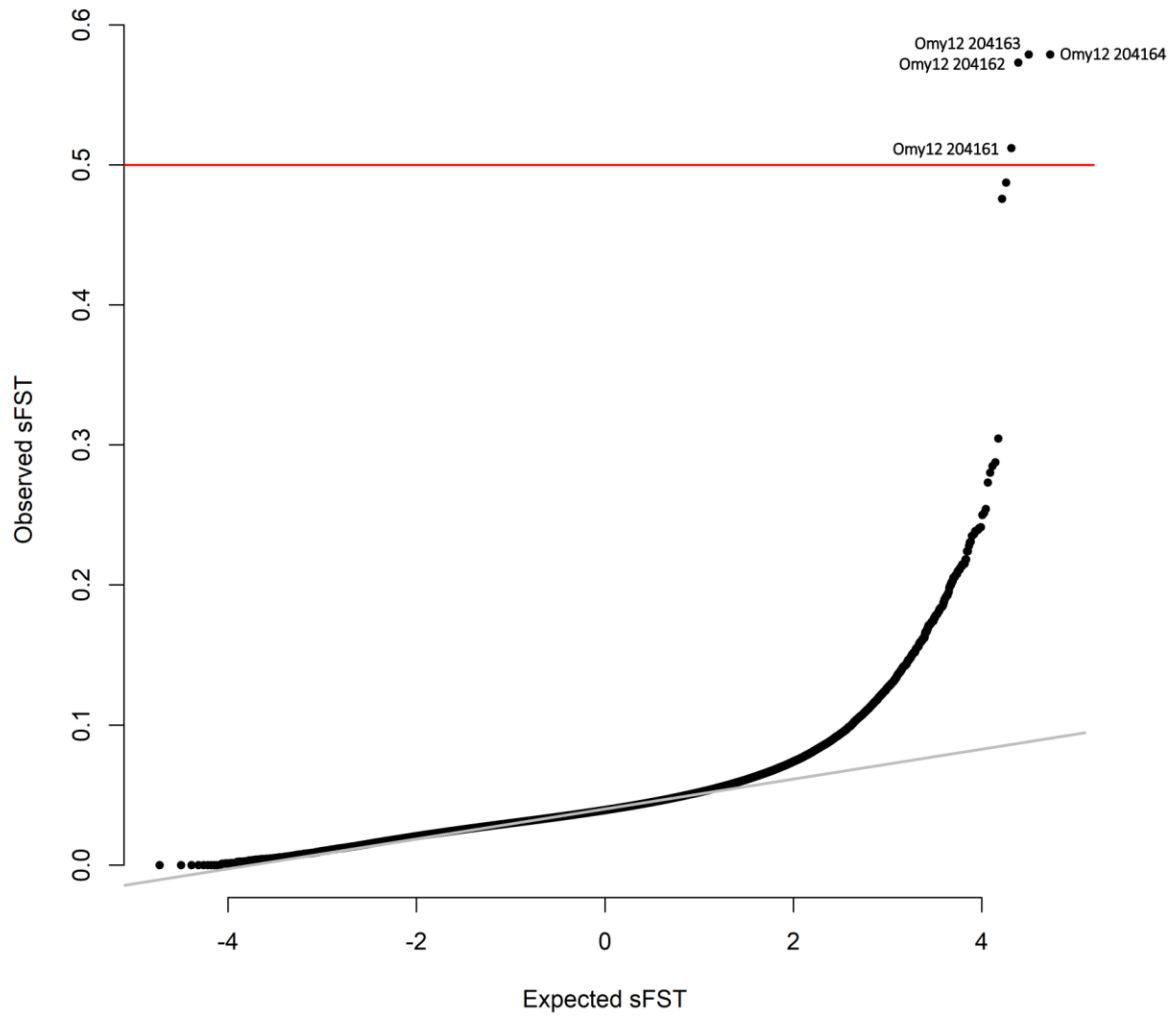
Supplementary Figure 3. Manhattan plot of FST scores. The minimum coverage threshold was 15 reads and the minor allele frequency was ≥ 0.05 .



Supplementary Figure 4. Number of overlapping loci and correlation of loci at different numbers of loci retained in Random Forest (RF) analysis.



Supplementary Figure 5. Local score Manhattan plot with male and female *O. mykiss* individuals included. The minimum coverage threshold was 15 reads and minor allele frequency was ≥ 0.05 .



Supplementary Figure 6. QQ-plot of observed compared to expected sliding window FST (sFST) data for all loci analyzed. The red line represents the threshold of significant sFST values and all loci surpassing the significance threshold are labelled.