OXFORD

Original Article

# Polygenic Basis and the Role of Genome Duplication in Adaptation to Similar Selective Environments

**Matthew A. Campbell,** Eric C. Anderson, John Carlos Garza, and Devon E. Pearse

From the Department of Ecology and Evolutionary Biology, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA (Campbell, Pearse); and the Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, 110 McAllister Way, Santa Cruz, CA 95060, USA (Campbell, Anderson, Garza, and Pearse).

Address correspondence to D. E. Pearse at the address above, or e-mail: Devon.Pearse@NOAA.gov.

## Abstract

Genetic changes underlying adaptation vary greatly in terms of complexity and, within the same species, genetic responses to similar selective pressures may or may not be the same. We examine both complex (supergene) and simple (SNP) genetic variants occurring in populations of rainbow trout (*Oncorhynchus mykiss*) independently isolated from ocean access and compared them to each other and to an anadromous below-barrier population representing their ancestral source to search for signatures of both parallel and nonparallel adaptation. All landlocked populations displayed an increased frequency of a large inversion on chromosome Omy05, while 3 of the 4 populations exhibited elevated frequencies of another inversion located on chromosome Omy20. In addition, we identified numerous regions outside these 2 inversions that also show significant shifts in allele frequencies consistent with adaptive evolution. However, there was little concordance among above-barrier populations in these specific genomic regions under selection. In part, the lack of concordance appears to arise from ancestral autopolyploidy in rainbow trout that provides duplicate genomic regions of similar functional composition for selection to act upon. Thus, while selection acting on landlocked populations universally favors the resident ecotype, outside of the major chromosomal inversions, the resulting genetic changes are largely distinct among populations. Our results indicate that selection on standing genetic variation is likely the primary mode of rapid adaptation, and that both supergene complexes and individual loci contribute to adaptive evolution, further highlighting the diversity of adaptive genomic variation involved in complex phenotypic evolution.

Understanding the process of adaptation, in which the phenotypes of individuals in a population respond to selection to increase fitness in the current environment, is a key line of inquiry in biology (Orr 2005). Heritable phenotypic change through genetic alterations is an important mechanism through which organisms can respond to new conditions, and the adaptive genetic variation that is selected for

in such cases can be identified using a variety of genetic techniques (Holderegger et al. 2006; Barrett and Hoekstra 2011; Hoban et al. 2016). What constitutes adaptive genetic variation varies greatly in terms of complexity. A simple example is that of the single nucleotide substitution causing a non-synonymous change that results in sickle cell anemia in humans (Pauling et al. 1949; Ingram 1956). Alternatively, supergenes—tightly linked clusters of genes that are inherited as a single unit—can contribute to adaptation (Fisher 1930; Schwander et al. 2014). Major life-history differences, local adaptation, and reproductive isolation have all been found to be associated with chromosomal inversion supergenes that capture multiple genes and suppress recombination in heterozygous individuals in diverse taxa (Kirkpatrick 2010; Lowry and Willis 2010; Kunte et al. 2014; Kirubakaran et al. 2016; Küpper et al. 2016; Lamichhaney et al. 2016).

The identification of genetic variation contributing to adaptation leads to more questions, such as the importance of existing supergene regions relative to adaptations involving single genes or de novo mutations, and the extent to which similar phenotypic selection results in parallel changes in the same genomic regions, representing genomic "islands" of divergence (Nosil and Feder 2012; Ruegg et al. 2014). Genomic islands of divergence have been implicated in adaptation in many species, notably with threespine stickleback (*Gasterosteus aculeatus*) (e.g., Hohenlohe et al. 2010; Marques et al. 2016) and sockeye salmon (Larson et al. 2017), and may originate from divergent selection. However, islands of genomic differentiation may also result not from divergent selection but due to variation in recombination rate, background selection, genetic drift, or other processes not related to adaptation or speciation (Noor and Bennett 2009; Cruickshank and Hahn 2014; Wolf and Ellegren 2017). Thus, a detailed characterization of the relationship between adaptive environments and genotypic changes, which may take on diverse forms, has implications for both evolutionary and conservation biology (Hansen et al. 2012; Pearse 2016).

Across much of their range, rainbow trout (*Oncorhynchus mykiss*) have been affected by human-built barriers to movement such as dams and road culverts (Sheer and Steel 2006; Clemento et al. 2009; Pearse and Campbell 2018; Winans et al. 2018). Many populations of landlocked rainbow trout have been created over the past century through the construction of barrier dams and transplantation into previously fish-less areas above natural waterfall barriers. Rainbow trout naturally occur as ecotypes with different migratory tendencies and timing; however, all ecotypes generally spawn in the winter or spring, interbreed when sympatric, and are iteroparous (Quinn 2005). Individual fish may be freshwater residents, not undertaking any large-scale migrations, or may migrate to the sea, returning to freshwater to spawn as large anadromous adults known as "steelhead" (Busby et al. 1996). Between these 2 extremes, fish may migrate within river and lake basins. In populations above dams and waterfalls, genetic contributions from anadromous fish are generally no longer possible, and any fish exhibiting migratory behavior are lost from the population by passing downstream over the barrier (Clemento et al. 2009; Pearse et al. 2009; Northcote 2010). As such, anadromy will be selected against, but the strength and action of such selection will vary depending on the conditions above the barriers (e.g., Pearse et al. 2014; Phillis et al. 2016; Leitwein et al. 2017).

Migratory phenotypes in rainbow trout have been demonstrated to have a genetic basis (e.g., Nichols et al. 2008; Hale et al. 2013; Hess et al. 2016). Therefore, when the relative reproductive success of migratory and nonmigratory individuals varies, the propensity of a population to produce migratory individuals will be affected, with corresponding changes in the frequency of genetic variants associated with migration. A single genomic region has previously been shown to be associated with resident and anadromous life history variation in rainbow trout (Pearse et al. 2014, 2019). This migration associated region (MAR; Leitwein et al. 2017) is a striking example of rapid changes in allele frequencies due to a sudden and sustained change in selective pressures. Recently, a new chromosome-scale genome assembly revealed that the MAR is a large chromosomal rearrangement consisting of a double inversion on Omy05 that links at least 4 separate homeologous regions originating from the ancestral whole genome duplication in salmonids (Ss4R) 88 mya into a single supergene (Pearse et al. 2019).

Characterization of the Omy05 rearrangement indicates that the ancestral form is associated with anadromy (A-type) and the rearranged form with residency (R-type), particularly in the southern part of the rainbow trout's natural range in the Eastern Pacific (Pearse et al. 2019). The Omy05 rearrangement is a large region, containing >1000 protein-coding genes of diverse functions, with sex-dependent dominance strongly influencing migration-related phenotypes. The frequency of Omy05 rearrangement haplotypes has been shown to be influenced by the relative reproductive success of migratory individuals in a given population, including in systems with restricted access to the ocean, but supporting adfluvial freshwater migrations to lakes or reservoirs (Pearse et al. 2014; Leitwein et al. 2017; Arostegui et al. 2019).

Despite the strong apparent association of the Omy05 rearrangement on migration, there is evidence that the effects of the Omy05 rearrangement are environmentally temperature-dependent, driving rapid early development and juvenile growth rate (e.g., Sundin et al. 2005; Nichols et al. 2007; Nichols et al. 2008; Hale et al. 2013; Rundio et al. 2020). However, there are also numerous other genomic regions associated with migration and associated traits (e.g., Nichols et al. 2008; Martínez et al. 2011; Hecht et al. 2012; Hale et al. 2013). Moreover, the Ss4R duplication has resulted in the presence of duplicates of many genes, providing diverse targets for selection to act on (Ohno 1970; Allendorf and Thorgaard 1984). This raises the intriguing possibility that the numerous genomic regions implicated in rainbow trout migratory behavior may partially be a result of alternatively selected genomic regions originating from the Ss4R. Here, we use RADseq data aligned to the recently improved assembly of the rainbow trout genome to investigate evidence for selection and parallel genomic adaptation in coastal California, and to identify the genomic architecture of adaptation in landlocked rainbow trout populations. We apply a comparative approach based on a novel analytical framework (Ruegg et al. 2014), and identify concordant and discordant patterns among multiple populations that have independently adapted to similar novel environmental conditions. These results improve our understanding of the evolution of adaptive genomic variation and touch upon major, long-standing questions in adaptation genetics (Orr 2005).

## Materials and Methods

### Study Populations

Details of sample collection have been described previously (Pearse et al. 2009; Abadía-Cardoso et al. 2016). For the present study, we utilized samples from 4 populations of rainbow trout that are now isolated above barriers to anadromy and a sample of below-barrier individuals representative of anadromous

**Table 1.** For each comparison of a landlocked population with the anadromous reference, the sample size, number of SNPs identified, and number of significantly elevated $F_{ST}$ RAD-locus-indices are given

| Comparison | Sample size of landlocked population | Number of SNPs called | Number of SNPs analyzed by swainysmoother | Number of total loci analyzed by swainysmoother | Significantly elevated $F_{ST}$ loci on Omy05 | F(R-type) Omy05 | Significantly elevated $F_{ST}$ loci on Omy20 | F(Type II) Omy20 | Other significantly elevated $F_{ST}$ loci | Other chromosomes with significantly elevated $F_{ST}$ loci |
|---|---|---|---|---|---|---|---|---|---|---|
| Big Creek resident | 12 | 41 447 | 17 221 | 8005 | 8 | 0.46 | 0 | 0.02 | 15 | 4, 8, 15, 17, 19, 26 |
| San Gabriel | 10 | 49 799 | 23 424 | 11 129 | 0 | 0.50 | 51 | 0.70 | 19 | 2, 9, 12, 16 |
| San Luis | 8 | 39 115 | 18 300 | 9016 | 31 | 0.88 | 1 | 0.32 | 20 | 2, 6, 12, 14, 17, 18, 24 |
| Matilija | 18 | 38 839 | 17 479 | 8638 | 113 | 0.75 | 0 | 0.39 | 27 | 3, 9, 10, 19, 25, 27, 28 |

The number of loci with significantly elevated $F_{ST}$ specific to Omy05 (all located in the inversion region) and Omy20 (41 of 52 located in the inversion region) are presented separately, with the frequencies of the Omy05 rearrangement (R-type), and the inversion polymorphism on Omy20, (Type II), indicated. F(R-type) from the Big Creek anadromous population is 0.18 and the F(Type II) is 0.04.

steelhead (Table 1). Due to sampling limitations in southern California where anadromous steelhead are rare and endangered as well as constraints on the number of samples that could be analyzed, we chose to focus on replicated instances of parallel adaptation of resident trout from a single sample representing a common steelhead ancestor. Steelhead populations throughout the coast of California exchange more migrants and have much higher connectivity than do isolated above-barrier populations (Clemento et al. 2009; Pearse et al. 2009). Thus, we chose to represent coastal steelhead with a single population sample, and compare it to 4 above-barrier populations for parallel and divergent changes from that common source over the past ~100 years since their isolation. Three of the landlocked populations are coastal California rainbow trout that are now isolated above dams in low-volume, high gradient tributary streams with limited potential for smaller scale, non-anadromous (adfluvial) migrations (Abadía-Cardoso et al. 2016). The San Gabriel River watershed is represented by samples from the West Fork above Cogswell Reservoir (WGS84 34.247°, −118.018°). The San Luis Rey River watershed further to the south was sampled in the West Fork above Lake Henshaw (WGS84 33.332°, −116.819°). The Ventura River watershed is represented by Matilija Creek above Matilija Reservoir (WGS84 34.506°, −119.384°). The fourth landlocked population is above a waterfall on Big Creek, within the Scott Creek drainage in Santa Cruz County (WGS84 37.092°, −122.209°) and was founded by transplants from below the waterfall approximately 100 years ago (Pearse et al. 2009). Thus, these 4 populations represent recent replicate isolations above barriers of fish that historically had access to the ocean; these above-barrier populations have significant differences (above waterfall vs. dams, distance from the common below-barrier population, access to reservoir habitat, etc.) but have a common history of adaptation to a similar resident rainbow trout ecotype. An anadromous steelhead population from Big Creek was used for comparison with all above-barrier rainbow trout populations, providing the opportunity to investigate parallel evolution to residency in similar environments. Characteristics of populations sampled in this study are further described through a map, phylogenetic relationships, and population genetic relationships in Supplementary Document S1.

## DNA Sequencing and Quality Control

Restriction site associated DNA sequencing (RADseq) libraries were generated following the protocol of Ali et al. (2016). Sequencing was conducted with 100 base pair (bp), paired-end sequencing. Demultiplexed sequences were trimmed for possible contamination with sequences from library generation and low quality bases with Trimmomatic version 0.36 (Bolger et al. 2014). Remaining paired-end reads were mapped to the recently released rainbow trout genome (NCBI GenBank assembly GCA 002163495.1; Pearse et al. 2019) with BWA version 0.7.5a-r405 using the mem algorithm (Li and Durbin 2009). The resulting Sequence Alignment/Map (SAM) files were converted to Binary Alignment/Map (BAM) files with SAMtools version 1.3 (Li et al. 2009). Samples with fewer than one million mapped reads were omitted from further analyses.

## Identification of Genomic Regions Under Differential or Parallel Selection

To identify regions of the genome under differential selection, we compared each of the 4 landlocked above-barrier populations to the Big Creek anadromous population, resulting in 4 pairwise comparisons. In each case, genotypes were called independently from sorted and indexed BAM files of the anadromous Big Creek and one landlocked population with ANGSD version 0.911- 51-g57d0264 (Korneliussen et al. 2014). Controlling for the quality of SNPs and for missing data was undertaken with the following settings: -doSaf 1 -doPost 2 -GL 2 -doMaf 2 -doMajorMinor 1 -minMaf 0.05 -minMapQ 30 -minQ 30 -SNP_pval 2e-3 -postCutoff 0.95, with an 85% missing data threshold (-minInd 22 with 26 individuals with greater than one million mapped reads, for example). The default ANGSD specification to include only properly paired reads was kept. The other settings affected the minimum minor allele frequency (-minMaf 0.05), the quality of bases and mapping quality to the genome (-minMapQ 30 -minQ 30) and the confidence in the called SNPs (-SNP_pval 2e-3 -postCutoff 0.95). For each BAM file, sites in the genome that had at least 10× coverage were identified through the generation of a pileup file with samtools (Li et al. 2009) and filtering with awk (Aho et al. 1988).

To identify genomic regions that are more diverged than the neutral expectation, and thus represent potential genomic islands

of divergence (Nosil and Feder 2012), the combination of a genotype file from ANGSD and coverage at each site was processed by a modified version of *swainysmoother* (Ruegg et al. 2014). The version of *swainysmoother* produced for this manuscript (available at https://github.com/MacCampbell/swainysmoother) was altered from original code (https://doi.org/10.5061/dryad.73gj4) to enable use with a reference genome, as well as to provide filtering of alignments to remove problematic regions based on coverage and length criteria. Regions of continuous sufficient coverage (>10×) at each restriction site may be from 100 up to 400 bp at each restriction site, and are referred to as loci in this article. *Swainysmoother* assigns a unique number to each locus and identifies those elevated in $F_{ST}$ by smoothing across a sliding window of 1 Mb centered on a site (see Ruegg et al. 2014 for details). Spatial autocorrelation was accounted for by generating 25 000 replicate smoothed lines, where, in each replicate, values of SNPs are permuted among locations on a chromosome. Thus, the replicates represent a scenario of no spatial autocorrelation along the chromosome, and loci with significantly elevated $F_{ST}$ exceed all 25 000 replicate smoothed values. This approach provides a very conservative identification of genomic regions with significantly elevated $F_{ST}$. Genomic regions likely under differential selection between each of the landlocked populations versus the anadromous population were identified by finding elevated-$F_{ST}$ loci in pairwise comparisons of the below-barrier sample and each above-barrier population.

Finally, signatures of likely parallel adaptation were identified through pairwise comparisons among the 4 landlocked populations (6 total comparisons), utilizing the same SNP calling and *swainysmoother* methods as described above. In contrast to the previous analysis, each pairwise comparison between 2 landlocked populations was made with the goal of identifying a "peak-and-valley" pattern generated by parallel selective sweeps on the same adaptive variants in populations with different genetic backgrounds as described by Roesti et al. (2014). Thus, we identified haplotype tracts with a central region of low $F_{ST}$, indicating parallel selection on the same variant, surrounded by elevated $F_{ST}$ peaks due to hitchhiking of different flanking sequences in each population.

### Identification of Omy05 and Omy20 Inversion Haplotypes

Previously, 4970 diagnostic variants were identified between A- and R-type Omy05 rearrangements (Supplementary Table S1; Pearse et al. 2019). These unique variants were filtered to 4108 biallelic SNPs that diagnose the Omy05 rearrangements to provide a reference data set to assign A- or R-type genotypes to individuals using a custom *perl* script (https://github.com/MacCampbell/scripts/blob/master/haplotyperOmy05.pl). Population level frequencies of the A- and R-type haplotypes were then estimated with standard error calculated as $\sqrt{\frac{p(1-p)}{2N}}$ with $p$ as the observed allele frequency and $N$ the sample size (Hartl and Clark 1997).

Genomic investigation of rainbow trout across its range also identified a chromosomal inversion on Omy20 without a known phenotypic association (Pearse et al. 2019). Output from *swainysmoother* indicated highly differentiated regions on this chromosome; however, diagnostic markers were not previously available to determine frequencies of this inversion. We defined a set of putatively diagnostic SNPs for the Omy20 inversion haplotype using the San Gabriel population as it appeared to have the highest frequency of an inversion haplotype (Type II) alternate to that of the Big Creek anadromous population. SNPs that were present in all San Gabriel

individuals in the genotypes produced for the *swainysmoother* analysis and were in complete linkage were identified as diagnostic. Sixty-nine SNPs satisfied these criteria (Supplementary Table S2) and genotypes from all populations were assessed in the genotype calls produced by ANGSD in each of the 3 remaining above-below comparisons. However, due to the effects of SNP calling in a pairwise fashion, it was unclear if some populations had both inversion haplotypes or if one haplotype occurred at a low frequency and was not genotyped. For example, in identifying Omy20 inversion variants from the Big Creek anadromous and Big Creek resident pairwise comparison SNP calls, the very low frequency of one Omy20 inversion variant led to very few of the diagnostic SNPs being present and ambiguity in identifying the Omy20 inversion haplotypes in the samples. Therefore, we examined genotype calls generated from all 5 sampling locations together with the same alignments and requirements as previously required for SNP calling between 2 populations to assess the presence of Omy20 inversion variants in the Big Creek above-waterfall population.

### Comparative Chromosomal Locations of Genomic Regions Associated With Migration

To allow direct comparison of the locations of previously identified genomic regions associated with migration on the same genomic coordinate system as this study, we used linked marker sequences to place them on the same genome assembly. Sequences of microsatellite loci identified as significantly differentiated by FDIST2 (Beaumont and Nichols 1996; Beaumont and Balding 2004) by Martínez et al. (2011), microsatellites or other loci in linkage groups from Nichols et al. (2008), and microsatellites and RAD loci from Hecht et al. (2012) that were significant genome-wide at the 95% threshold were downloaded from GenBank or obtained from supplemental material of the source paper. Each sequence was searched against the rainbow trout genome with BLASTN (Altschul et al. 1997; Morgulis et al. 2008) driven by a perl script (available at https://github.com/MacCampbell/scripts/blob/master/locateSeqs.pl). An expect value (*e* value) of 0.001 was specified. Hits were further filtered by requiring 95% sequence identity and 95% alignment length of queried sequence to target sequence. It was possible for a marker to align more than once, though only the best match was retained.

Finally, as homologous blocks originating from a WGD are similar in composition, these regions may offer alternative selective targets in similar environments. We tabulated high $F_{ST}$ loci from different pairwise comparisons between the below-barrier and above-barrier populations that occurred inside or outside of the 88 known pairs of homeologous blocks from the Ss4R mapped in the rainbow trout genome (Pearse et al. 2019).

## Results

### DNA Sequencing and Sequence Data Quality Control

After sequence quality control, the data set contained 8–18 samples per population with greater than one million reads mapped (Table 1, Supplementary Table S3).

### Identification of Genomic Regions Under Differential or Parallel Selection

The number of individuals and SNPs included in analyses for differential selection are indicated in Table 1. The number of SNPs analyzed

for each pair of populations ranged from ~17000 to 24000 (Table 1). From 23 to 140 loci were identified in each pairwise above- vs. below-barrier comparison as having an elevated $F_{ST}$, with a total of 285 high $F_{ST}$ loci (Figure 1, Table 1). Among all pairwise population comparisons, all elevated $F_{ST}$ loci were unique in position. Multiple loci on chromosomes Omy05 and Omy20 with known inversions were clearly identified by *swainysmoother* (Figure 2, Table 1). All elevated $F_{ST}$ loci on chromosome Omy05 were inside the inversion, while 40/51 of such loci found on Omy20 in the Big Creek below-barrier to San Gabriel comparison were within the inversion on that chromosome. On other chromosomes, 15–27 loci per population were indicated to be elevated in $F_{ST}$, for a total of 81 additional loci located on 19 chromosomes identified across all above- to below-barrier comparisons (Table 1).

Among the pairwise comparisons between landlocked populations, the number of SNPs identified for each of the 6 comparisons ranged from 10 thousand to 36 thousand. Following the logic of

Roesti et al. (2014), 5 genomic regions showed the peak-and-valley signature of parallel selection (Figure 3, Table 2). The 5 regions showing evidence of parallel selection were found in 3 different pairwise population comparisons involving 4 separate chromosomes, including 2 distinct genomic regions on Omy01 found in the Big Creek-San Gabriel and San Luis-San Gabriel comparisons, respectively. Two additional regions on Omy03 and Omy12 were identified in the San Gabriel and San Luis comparison as potentially under parallel selection (Figure 1), whereas a region on Omy09 was similarly identified in the San Luis-Matilija comparison (Figure 3, Table 2).

## Identification of Omy05 and Omy20 Inversion Haplotypes

The frequencies of the R-type Omy05 MAR haplotype in all populations is presented in Table 1 and Figure 4A. As is typical for coast
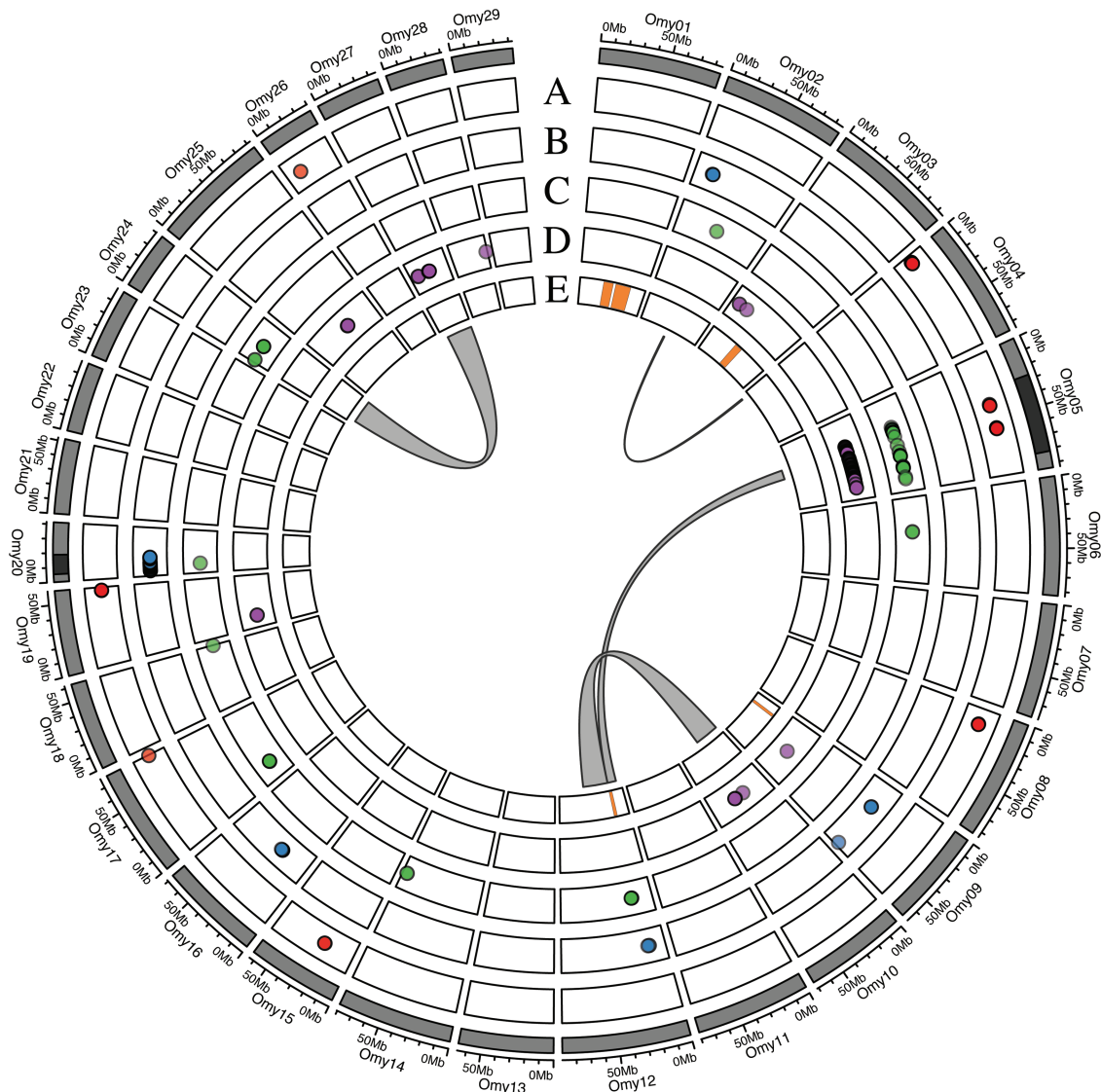


**Figure 1.** Circular representation of the rainbow trout genome with locations of high $F_{ST}$ loci indicated, as well as regions with evidence of parallel selection and homeologous pairs. Chromosomes are named and length in Mbp shown. Known inversion regions on Omy05 and Omy20 are shaded in the outer ring. For each of the pairwise comparisons, the locations of high $F_{ST}$ loci are shown with colored points: Track A, Big Creek Resident; Track B, San Gabriel; Track C, San Luis; Track D, Matilija; Track E, combined results from all *above-barrier to above-barrier* comparisons showing regions that may be under parallel selection (rectangles). Internal gray connectors show pairs of homeologous regions originating from the Ss4R that both contain high $F_{ST}$ loci. See online version for color figure.

## 2A



Position Along Omy05

## 2B



Position Along Omy20
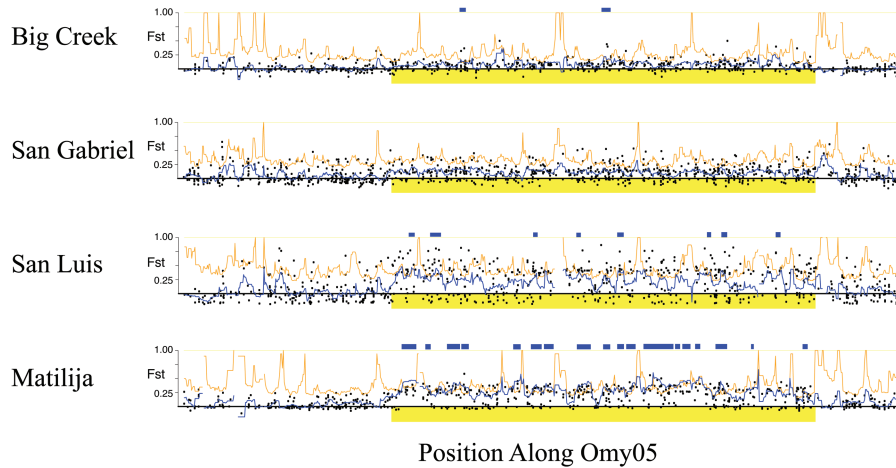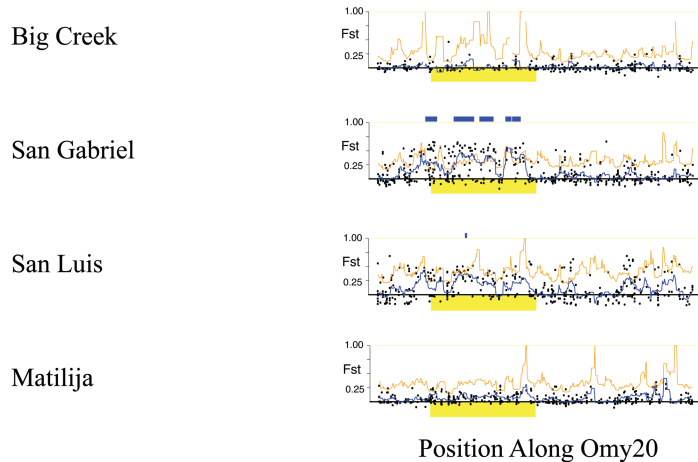
**Figure 2**. Output from *swainysmoother* for chromosome Omy05 (**A**) and chromosome Omy20 (**B**) from comparisons between the below-barrier and 4 above-barrier populations. The *x* axis is the position on the chromosome, the *y* axis is $F_{ST}$. For each plot, black points represent observed $F_{ST}$ at a site, the dark grey line the smoothed $F_{ST}$ centered on the corresponding site in the genome, and the light grey line the maximal $F_{ST}$ generated after 25 000 permutations of *swainysmoother* at the corresponding site. For each chromosome, the extent of the inversion region is indicated by the shaded rectangle below the x-axis. Regions where the observed $F_{ST}$ is greater than the maximal $F_{ST}$ generated through the permutations are indicated by dark blocks at the top of each chromosome plot. When the nucleotide diversity of pooled samples is relatively small at a site or when nucleotide diversity for one population is much smaller than the other population, these cases this may result in a negative $F_{ST}$ value that can be interpreted as 0 (Ruegg et al. 2014). See online version for color figure.

steelhead populations (Pearse et al. 2014), the Big Creek anadromous population a low frequency of this rearranged haplotype [F(R-type) = 0.18], while all landlocked populations had frequencies of ≥0.46. The frequency of the Type II Omy20 inversion variant is presented in Table 1 and Figure 4B. Both the Big Creek above- and below-barrier populations have very low frequencies of the Type II inversion, with only a single heterozygous individual in each population. All other above-barrier populations had much higher frequencies of the Type II Omy20 inversion polymorphism in comparison (Table 1 and Figure 4).

### Comparative Chromosomal Locations of Genomic Regions Associated With Migration

To compare our results with those of previous studies, we aligned the positions of markers associated with migration in those studies to the updated *O. mykiss* genome assembly (Pearse et al. 2019;

Supplementary Table S4). Twelve of the 14 microsatellites from Martínez et al. (2011) were placed on 10 chromosomes (5, 6, 10, 15, 18, 19, 21, 22, 23, 29) (Supplementary Table S4). Eleven linkage groups (Nichols et al. 2008) were placed onto 11 chromosomes (1, 3, 5, 8, 19, 14, 16, 19, 20, 23, 28). Nine quantitative trait loci (QTL) from (Hecht et al. 2012) are placed onto 7 chromosomes (1, 5, 8, 11, 12, 20, 28). A subset of *swainysmoother* output is included in Supplementary Table S4 to ease comparisons between this manuscript and others (See "Adaptation to Residency in Rainbow Trout" in Discussion).

### Evidence for Differential Selection on Paired Homeologous Blocks

Overall, nearly one-quarter (23%, 66/285) of the high $F_{ST}$ loci identified in below-barrier to above-barrier comparisons were located within 4 pairs of homeologous blocks originating from the salmonid-specific whole-genome duplication (Figure 1, Supplementary Table
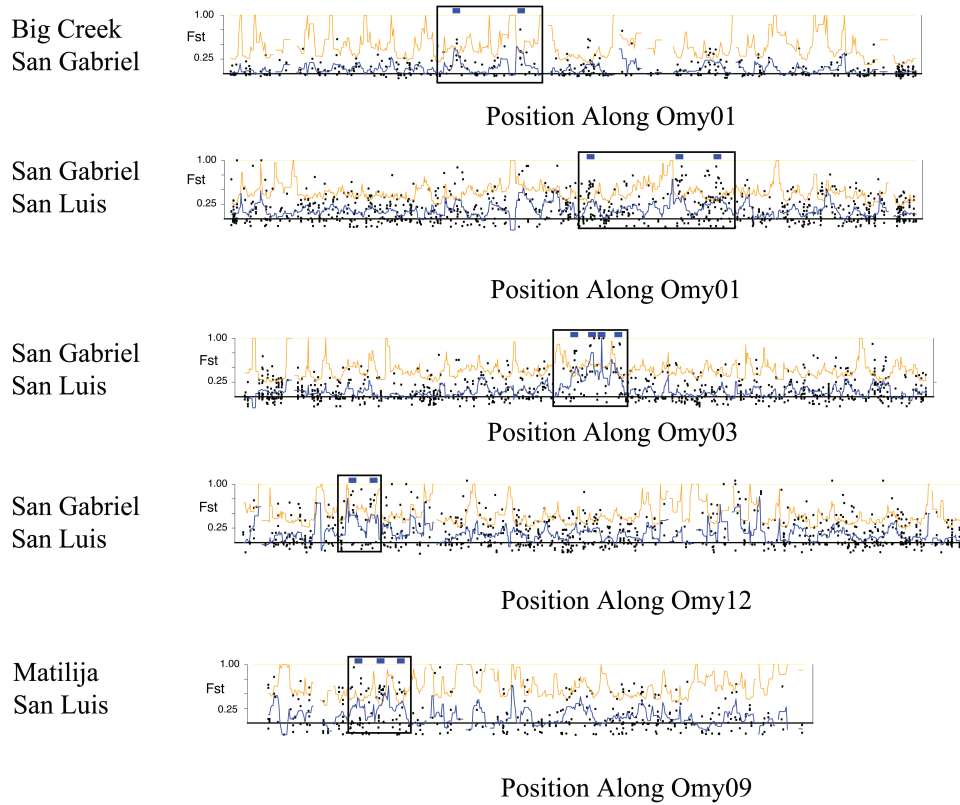
**Figure 3.** Output from *swainysmoother* for the 5 pairwise comparisons that indicated parallel adaptation between above-barrier populations. Regions fitting the patterns described by Roesti et al. (2014) for identifying parallel adaptation are indicated within black boxes. The 2 populations involved are indicated, with the *x* axis representing the position on the chromosome, and the *y* axis $F_{ST}$. Black points represent observed $F_{ST}$ at a site, the dark grey line the smoothed $F_{ST}$ centered on the corresponding site in the genome, and the light grey line the maximal $F_{ST}$ generated after 25 000 permutations of *swainysmoother* at the corresponding site. Significantly high $F_{ST}$ peaks are indicated by dark blocks at the top of each chromosome plot. The elevated $F_{ST}$ peaks bounding non-elevated $F_{ST}$ valleys are listed in Table 2. When the nucleotide diversity of pooled samples is relatively small at a site or when nucleotide diversity for one population is much smaller than the other population, these cases may result in a negative $F_{ST}$ value that can be interpreted as 0 (Ruegg et al. 2014). See online version for color figure.

**Table 2.** Low $F_{ST}$ regions surrounded by high $F_{ST}$ peaks identified from pairwise comparisons among landlocked populations

| Comparison | Chromosome | High $F_{ST}$ loci positions (Mbp) |
|---|---|---|
| Big Creek resident to San Gabriel | 1 | 30.6, 31.2, 42.0, 42.7, 43.0 |
| San Gabriel to San Luis | 1 | 46.3, 46.5, 60.0, 65.4, 65.6, 65.7, 65.8 |
| San Gabriel to San Luis | 3 | 36.6, 36.7, 36.8, 36.9, 37.1, 37.3, 40.6, 40.7, 40.8, 41.1, 41.2, 46.1, 46.4 |
| San Gabriel to San Luis | 12 | 15.4, 15.7, 16.0, 16.1, 16.4, 19.1, 19.3, 19.5 |
| San Luis to Matilija | 9 | 11.7, 13.3, 15.4 |

For comparisons with evidence of parallel adaptation, the chromosome and range of positions are shown, with high $F_{ST}$ loci rounded to the nearest 0.1 of a Mbp. Multiple loci may follow sequentially after each position but are omitted for clarity.

S5). However, as chromosomal inversions result in strong linkage disequilibrium, the loci within the 2 known inversions are not independent. Taking into account the non-independencies of loci located in chromosomal inversions results in a similar estimate of the proportion of the loci located within these 4 homeologous block pairs (25%, 23/94). Interestingly, 3 of the homeologous pairs involve 2 of the below-barrier to above-barrier comparisons, including high $F_{ST}$ indices on Omy24 present in the San Luis comparison pairing with high $F_{ST}$ indices on Omy27 in the Matilija comparison. In addition, a section of Omy05p (33–47 Mbp) found to contain high $F_{ST}$ RAD-locus-indices in the Big Creek, San Luis and Matilija comparisons corresponds to a section of Omy12p with high $F_{ST}$ RAD-locus-indices also found in the San Gabriel comparison. However, no high

$F_{ST}$ RAD-locus-indices are observed on Omy05 in the Big Creek below-barrier to San Gabriel comparison.

## Discussion

We used genome-wide analysis to investigate patterns of genomic association in multiple populations of rainbow trout adapting to a resident life-history phenotype from a common anadromous ancestors. We found a complex polygenic pattern of shared and unique genomic adaptations in different trout populations. While the major-effect supergenes on Omy05 and Omy20 were consistently associated with this transition, most other differentiated genomic regions were not shared among different resident trout populations when

## 4A



## 4B
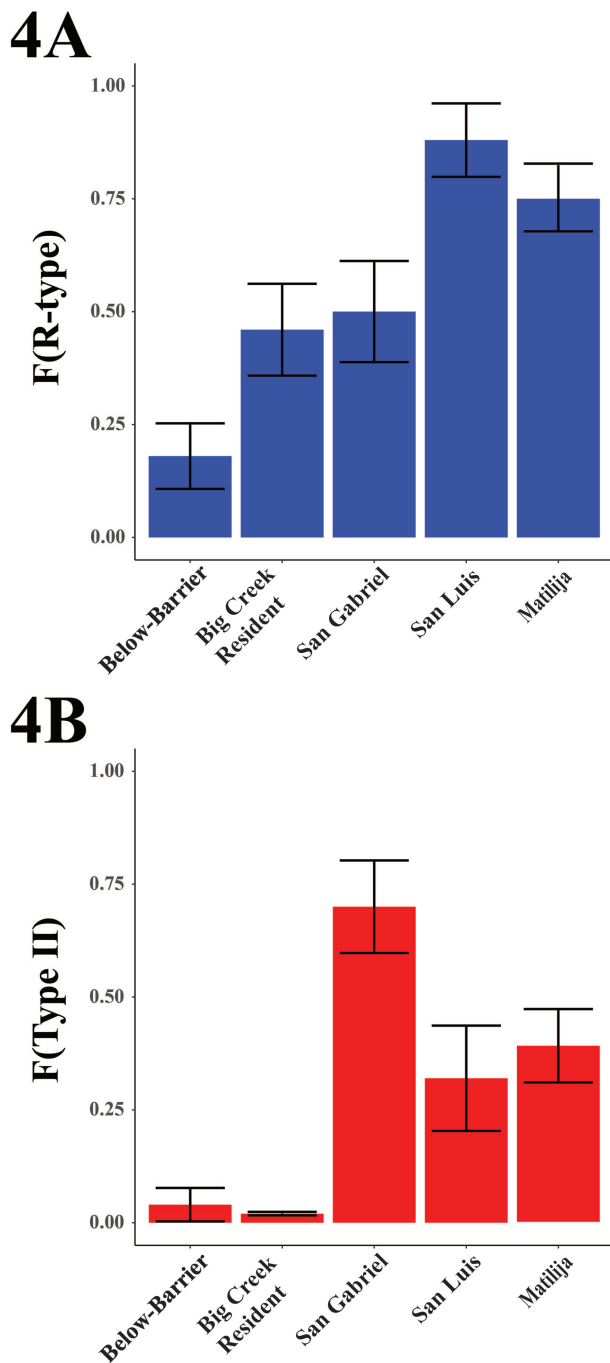


**Figure 4.** (**A**) The frequency of the Omy05 rearrangement associated with residency (R-type), and (**B**) the predominant Omy20 inversion haplotype found in land-locked populations (Type II). For each inversion frequency measure, standard error (SE) is shown.

compared to a common anadromous ancestor. The ancestral whole-genome duplication in salmonid fishes produced homeologous regions that may provide increased material for nonparallel adaptation.

### Detection of Genomic Regions Under Selection

Methods to identify significantly differentiated genomic regions were originally developed for anonymous markers without knowledge of their relative locations or placement in the genome (e.g., Arlequin, BayeScan, and FDIST2; Beaumont and Nichols 1996; Foll

and Gaggiotti 2008; Excoffier et al. 2009). Each marker locus is therefore treated independently in these outlier detection methods. With few loci available and the genomic coordinates unknown, this approach is reasonable. However, the scale of genetic data collection has grown substantially, such that the identification of large numbers of variants is routine and the number of available genome assemblies is large and expanding, offering the potential to greatly increase our power to identify patterns associated with selection on adaptive genomic variation. For example, when we examined our data from the Big Creek population pair with BayeScan, using the same SNP calling as used in the paper with *swainysmoother*, we detected only one outlier SNP (on Omy15 at site 22,106,770), and no SNPs from the Omy05 rearrangement were detected as outliers despite a strong a priori expectation that this region should be differentiated between the above- and below-barrier populations from Big Creek. In contrast, using *swainysmoother*—a method that incorporates genomic SNP density and background $F_{ST}$ differences to stringently identify regions of elevated $F_{ST}$—we identified 5 significantly elevated loci from the Omy05 MAR and numerous other genomic regions located on 6 other chromosomes with the same data from the Big Creek population pair (Figures 1 and 2). These included 2 high $F_{ST}$ loci on Omy15 in the Big Creek comparison within 2 Mbp of the outlier SNP indicated by BayeScan, indicating some congruence between the 2 approaches.

The properties and processes that cause *swainysmoother* and other methods to recognize an island of genomic divergence are likely varied, and the influence of recombination, genetic drift, background selection, genetic draft, and other processes in creating high $F_{ST}$ loci should be considered (Noor and Bennett 2009; Cruickshank and Hahn 2014; Wolf and Ellegren 2017). In particular, suppressed recombination occurs with chromosomal inversions (Kirkpatrick 2010), and the method may be especially effective at identifying genomic regions with low recombination. The recent isolation of the above-barrier populations likely indicates that these genomic islands are not a result of drift, as lineage-sorting from drift is a slow-acting process relative to selection, even in small populations (Cruickshank and Hahn 2014). Until the *swainysmoother* method is more thoroughly evaluated, our conclusion is that we have identified regions that may be under local selection in rainbow trout. Therefore, utilizing the new rainbow trout assembly and data from other studies to place the high $F_{ST}$ loci identified into context with other sources of evidence (e.g., quantitative trait loci) can lead to more conclusive results supported by independent lines of evidence. However, it is also worth noting that populations examined here may have been recently bottlenecked, potentially increasing the difficulty of identifying adaptive loci (Poh et al. 2014).

### Adaptation to Residency in Rainbow Trout

Numerous chromosomes without known inversions have been implicated as important for adaptation to residency through migration- and smoltification-related traits. Positions on these chromosomes of markers between studies may be closely located, but not necessarily. For example, the locations of significant markers on Omy10 in both Martínez et al. (2011) and Nichols et al. (2008) are close, with approximately 1.3 megabase pairs (Mbp) between the microsatellite markers OMM1120 and OMM1050. However, no candidate loci on chromosome Omy10 were identified in Hecht et al. (2012). Furthermore, while several other chromosomes containing important regions have been identified in 2 or more studies, all such significant regions are actually quite distant from each other despite occurring on the

same chromosome (e.g., Omy08, Omy19, Omy20, and Omy28; Supplementary Table S4). These findings highlight the importance of variation in linkage disequilibrium when comparing signals of selection with linked loci.

Multiple quantitative trait loci also have been shown to co-localize on chromosomes Omy05 and Omy10 (Nichols et al. 2008). These particular chromosomes were identified in the present study as containing high $F_{ST}$ genomic regions in some, though not all, below-barrier to above-barrier comparisons. However, in the above-barrier to above-barrier population comparisons using the approach described by Roesti et al. (2014) for identifying parallel adaptation, we also identified 2 other genomic regions fitting these patterns that correspond to regions previously associated with migration related traits. First, in the Big Creek and San Gabriel pairwise comparison, the marker placing the KFact905 (body condition factor) QTL from Hecht et al. (2012) on Omy01 at 31.8 Mbp is surrounded by high $F_{ST}$ loci identified by *swainysmoother* around 30 and 42 Mbp. Similarly, the San Gabriel to San Luis comparison has high $F_{ST}$ loci near 46 and 60 Mbp on Omy01 that surround the microsatellite marker OMM1146 from linkage group OC6 (binary smolt phenotype) identified in Nichols et al. (2008) and located on Omy01 at 48.9 Mbp. Thus, our results add to a growing body of evidence that show that, in addition to the Omy05 and Omy20 supergenes, many separate genomic regions underlie adaptation to residency across populations (Nichols et al. 2008; Martínez et al. 2011; Hecht et al. 2012; Hale et al. 2013). Strikingly, however, the same genomic coordinates were not identified with high $F_{ST}$ loci in the different pairwise comparison, outside of the 2 primary chromosomal inversions. Thus, our results indicate that multiple genes and genomic regions of both large and small effect underpin adaptation of migration phenotypes in above-barrier rainbow trout populations.

## Big Creek Population Pair

Outlier loci in the Big Creek population pair were first investigated by Martínez et al. (2011), in a genome screen using 298 microsatellite loci to search for regions under differential selection. The most conservative approach used by Martínez et al. (2011) identified 14 outlier loci located on 11 chromosomes. Here, with data from 10s of thousands of loci and the methodology of *swainysmoother*, only 23 loci were identified as outliers elevated in $F_{ST}$. *Swainysmoother* is a conservative method that only identifies a locus to have elevated $F_{ST}$ when it is higher than all 25 000 permutations, so also has a low false positive rate. We identified some of the same chromosomes (Omy05, Omy15, Omy19) as Martínez et al. (2011) in the Big Creek comparison. However, given the relatively low genomic coverage of both the microsatellite and RADseq datasets, neither study has high power to resolve individual loci under selection (Lowry et al. 2017), and extensive concordance between the 2 studies was thus not expected. Nonetheless, RAD-locus-index 3967 was identified as elevated in $F_{ST}$ in this study, and is located 1.3 Mbp from the microsatellite locus BHMS426 identified by Martínez et al. (2011) in the large inversion region on Omy05 and are therefore tightly linked (Pearse et al. 2019).

## Insights Into Adaptation Genetics

Recent studies have made major advances in understanding the genomic basis of adaptation (e.g., Natarajan et al. 2016; Reid et al. 2016). However, the generality of such findings is not yet known, and many questions remain. The analyses presented here provide further insight into some of these questions.

## Are Supergenes Major Players in Most Adaptations?

Evidence of supergenes as broadly important in adaptation is increasing (Kunte et al. 2014; Küpper et al. 2016; Lamichhaney et al. 2016; Tuttle et al. 2016; Reviewed by Wellenreuther and Bernatchez 2018). Our results further highlight the role of the inversion supergene region on chromosome Omy05 in migration-related traits in *O. mykiss* (Pearse et al. 2014, 2019). High $F_{ST}$ loci from Omy05 contributed over half the total number of high $F_{ST}$ loci in this study (152/285 = 53%, Figure 2A) in comparisons between the below-barrier and above-barrier fish, providing additional data on the role of the Omy05 inversion supergene region on migration-related traits in *O. mykiss* (Pearse et al. 2014, 2019). In addition, we detected apparent differential selection on a smaller inversion located on chromosome Omy20 with 2 populations containing elevated $F_{ST}$ loci within this inversion (Figure 2B). This inversion also contributed substantially to the total number of high $F_{ST}$ loci (41/285 = 14%). Chromosome Omy20 was identified by both Nichols et al. (2008) and Hecht et al. (2012) as related to rainbow trout migration. The region of Omy20 with elevated $F_{ST}$ in the San Gabriel comparison overlaps with locus R43574 from Hecht et al. (2012). Given that very little is known about the distribution of the Omy20 inversion throughout the range of rainbow trout and that any phenotypic effects of Omy20 inversion haplotypes are unknown, our results provide evidence that there may be a relative increase in the frequency of Omy20 in the southern part of the rainbow trout range and/or that it may be under selection in above-barrier populations.

## Does Evolution of the Same Phenotype Involve the Same Genetic Basis?

All landlocked populations in this study are in environments where selection against anadromy and migration should be strong. While the same resident phenotype is being selected for, does this adaptation consistently involve selection in the same genetic regions? The overall increase in Omy05 R-type inversion haplotype frequencies is clearly shared among above-barrier populations as has been previously established (Pearse et al. 2014). From the 6 pairwise comparisons, we identified 5 genomic regions in 3 comparisons that showed evidence of parallel evolution. However, there is very little agreement among our comparisons in which genomic regions are underlying adaptation to a resident phenotype. Our results indicate that while the overall phenotypic response to a given novel environment may be predictably the same, populations are likely to arrive at the same phenotype through different genetic changes, possibly affecting the same molecular pathways. Similar results have been found in an examination of more than 50 bird species in which hemoglobin binding affinity increases were predicted and found in high altitudes, but the observed phenotypic change was generally not caused by parallel changes at key amino acids within that protein's DNA sequence (Natarajan et al. 2016). A study of rough periwinkle (*Littorina saxatilis*) similarly identified very few shared outlier loci (2–9% of total RADseq loci) despite common recent ancestry of ecotypes and parallel phenotypic divergence among 3 populations (Ravinet et al. 2016). Multiple pairs of freshwater and anadromous Atlantic salmon also shared few outlier loci with nonparallel divergence at numerous loci presumably of small effect (Perrier et al. 2013). Therefore, our findings are consistent with previous demonstrations of parallel phenotypic evolution involving few of the same, and many alternative, genetic changes. In contrast, however, recent findings on another key life-history trait, migration timing, in steelhead and Chinook salmon (*O. tshawytscha*), have clearly demonstrated a single-locus genetic basis for this complex trait (Hess et al. 2016; Prince et al. 2017; Thompson et al. 2020).

### How Do Duplicated Regions Contribute to Parallel Adaptive Processes?

Whole-genome duplications are found throughout the eukaryotic tree of life at both recent and ancient timescales, including within the common ancestor of all modern salmonids. This duplication occurred approximately 88 million years ago and is termed the Ss4R (Allendorf and Thorgaard 1984; Berthelot et al. 2014; Macqueen and Johnston 2014; Campbell et al. 2016; Van de Peer et al. 2017). The Ss4R has resulted in large duplicated regions throughout salmonid genomes (Allendorf and Thorgaard 1984; Berthelot et al. 2014; Lien et al. 2016). Genetic redundancy in such duplicated regions is hypothesized to facilitate the derivation of genes with new functions (Ohno 1970). However, gene copies arising from a whole-genome duplication—ohnologs—may take on subsets of their original functions or be conserved in function (Force et al. 1999; Rastogi and Liberles 2005; Sémon and Wolfe 2008; Warren et al. 2014; Campbell et al. 2019).

The highly polygenic basis of genomic adaptation implicated in rainbow trout life-history variation may be a result of duplicated regions originating from the Ss4R being alternately selected upon during parallel adaptation. Consistent with this hypothesis, we identified 4 pairs of homeologous blocks in alternative above-barrier to below-barrier comparisons, indicating potential differential selection is occurring on similar, but not identical, genomic regions (Figure 1, Supplementary Table S5). These 4 regions contain at least 23% of the high $F_{ST}$ loci identified in the study, pointing to the key role that whole-genome duplication may play in producing important adaptive genetic variation. While 3 of the homeologous blocks contain regions have not previously been implicated in adaptation to residency or anadromy, one pair of homologous blocks has one region in the Omy05 inversion and the other on Omy12, which was notably found to have 7 quantitative trait loci of importance in a previous study (Hecht et al. 2012).

The Omy05 inversion haplotype is broadly associated with residency, but there were no high $F_{ST}$ loci within the Omy05 inversion in the comparison of the San Gabriel to the below-barrier (Big Creek) population. However, there were high $F_{ST}$ loci in the corresponding homeologous region on Omy12 in that comparison (Figure 1, $n = 3$). Our finding that just 4 homeologous regions contain a substantial proportion of high $F_{ST}$ loci supports the view that redundant gene complements from whole-genome duplications are important for adaptive potential (Van de Peer et al. 2017), and is consistent with recent findings of apparent parallel selection on duplicated copies of the same genomic regions in other salmonids (Salisbury et al. 2020).

The landlocked populations examined in this study show genetic signals consistent with absence of reproductive success of anadromous individuals and strong selection against migration. This selection against migration affects not only the same genomic regions within populations, but also numerous separate genomic regions. Further investigation may indicate if the separate genomic regions that may be under selection in above-barrier populations actually form parts of the same molecular pathways contributing to adaptation to residency.

### Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

### Funding

### Acknowledgments

### Data Availability

Demultiplexed RADseq data, genotype calls, and *swainysmoother* output are available from the Dryad repository (https://doi.org/10.25338/B8KH0X). All code associated with analysis is available at GitHub (https://github.com/MacCampbell/swainysmoother and https://github.com/MacCampbell/scripts) or in the Dryad repository for this paper.

### References

Abadía-Cardoso A, Pearse DE, Jacobson S, Marshall J, Dalrymple D, Kawasaki F, Ruiz-Campos G, Garza JC. 2016. Population genetic structure and ancestry of steelhead/rainbow trout (*Oncorhynchus mykiss*) at the extreme southern edge of their range in North America. *Conserv Genet*. 17:675–689.

Aho AV, Kernighan BW, Weinberger PJ. 1988. *The AWK programming language*. Reading, MA: Addison-Wesley Longman Publishing Co., Inc.

Ali OA, O'Rourke SM, Amish SJ, Meek MH, Luikart G, Jeffres C, Miller MR. 2016. RAD capture (Rapture): flexible and efficient sequence-based genotyping. *Genetics*. 202:389–400.

Allendorf FW, Thorgaard GH. 1984. Tetraploidy and the evolution of salmonid fishes. In: Turner BJ, editor. *Evolutionary genetics of fishes* (Monographs in Evolutionary Biology). Springer US. p. 1–53.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25:3389–3402.

Arostegui MC, Quinn TP, Seeb LW, Seeb JE, McKinney GJ. 2019. Retention of a chromosomal inversion from an anadromous ancestor provides the genetic basis for alternative freshwater ecotypes in rainbow trout. *Mol Ecol*. 28:1412–1427.

Barrett RD, Hoekstra HE. 2011. Molecular spandrels: tests of adaptation at the genetic level. *Nat Rev Genet*. 12:767–780.

Beaumont MA, Balding DJ. 2004. Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol*. 13:969–980.

Beaumont MA, Nichols RA. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proc Biol Sci*. 263:1619.

Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, Bento P, Da Silva C, Labadie K, Alberti A, *et al.* 2014. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nat Commun*. 5:3657.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114–2120.

Busby PJ, Wainwright TC, Bryant GJ, Lierheimer LJ, Waples RS, Waknitz FW, Lagomarsino IV. 1996. *Status review of west coast steelhead from Washington, Idaho, Oregon, and California*. Springfield, VA: US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service.

Campbell MA, Ganley AR, Gabaldón T, Cox MP. 2016. The case of the missing ancient fungal polyploids. *Am Nat*. 188:602–614.

Campbell MA, Hale MC, McKinney GJ, Nichols KM, Pearse DE. 2019. Long-term conservation of ohnologs through partial tetrasomy following whole-genome duplication in Salmonidae. *G3 (Bethesda)*. 9:2017–2028.

Clemento AJ, Anderson EC, Boughton D, Girman D, Garza JC. 2009. Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. *Conserv Genet*. 10:1321.

Cruickshank TE, Hahn MW. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol*. 23:3133–3157.

Excoffier L, Hofer T, Foll M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity (Edinb)*. 103:285–298.

Fisher RA. 1930. *The genetical theory of natural selection*. Oxford, UK: The Clarendon Press.

Foll M, Gaggiotti O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*. 180:977–993.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 151:1531–1545.

Hale MC, Thrower FP, Berntson EA, Miller MR, Nichols KM. 2013. Evaluating adaptive divergence between migratory and nonmigratory ecotypes of a salmonid fish, *Oncorhynchus mykiss*. *G3 (Bethesda)*. 3:1273–1285.

Hansen MM, Olivieri I, Waller DM, Nielsen EE; GeM Working Group. 2012. Monitoring adaptive genetic responses to environmental change. *Mol Ecol*. 21:1311–1329.

Hartl DL, Clark AG. 1997. *Principles of population genetics*. 3rd ed. Sunderland (MA): Sinauer Associates, Inc.

Hecht BC, Thrower FP, Hale MC, Miller MR, Nichols KM. 2012. Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. *G3 (Bethesda)*. 2:1113–1127.

Hess JE, Zendt JS, Matala AR, Narum SR. 2016. Genetic basis of adult migration timing in anadromous steelhead discovered through multivariate association testing. *Proc Biol Sci*. 283:20153064.

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am Nat*. 188:379–397.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet*. 6:e1000862.

Holderegger R, Kamm U, Gugerli F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landsc Ecol*. 21:797–807.

Ingram VM. 1956. A specific chemical difference between the globins of normal human and sickle-cell anaemia haemoglobin. *Nature*. 178:792–794.

Kirkpatrick M. 2010. How and why chromosome inversions evolve. *PLoS Biol*. 8:e1000501.

Kirubakaran TG, Grove H, Kent MP, Sandve SR, Baranski M, Nome T, De Rosa MC, Righino B, Johansen T, Otterå H, *et al.* 2016. Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Mol Ecol*. 25:2130–2143.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*. 15:356.

Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. 2014. Doublesex is a mimicry supergene. *Nature*. 507:229–232.

Küpper C, Stocks M, Risse JE, Dos Remedios N, Farrell LL, McRae SB, Morgan TC, Karlionova N, Pinchuk P, Verkuil YI, *et al.* 2016. A supergene determines highly divergent male reproductive morphs in the ruff. *Nat Genet*. 48:79–83.

Lamichhaney S, Fan G, Widemo F, Gunnarsson U, Thalmann DS, Hoeppner MP, Kerje S, Gustafson U, Shi C, Zhang H, *et al.* 2016. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat Genet*. 48:84–88.

Larson WA, Limborg MT, McKinney GJ, Schindler DE, Seeb JE, Seeb LW. 2017. Genomic islands of divergence linked to ecotypic variation in sockeye salmon. *Mol Ecol*. 26:554–570.

Leitwein M, Garza JC, Pearse DE. 2017. Ancestry and adaptive evolution of anadromous, resident, and adfluvial rainbow trout (*Oncorhynchus mykiss*) in the San Francisco bay area: application of adaptive genomic variation to conservation in a highly impacted landscape. *Evol Appl*. 10:56–67.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup.

2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 25:2078–2079.

Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, Nome T, Hvidsten TR, Leong JS, Minkley DR, Zimin A, *et al.* 2016. The Atlantic salmon genome provides insights into rediploidization. *Nature*. 533:200–205.

Lowry DB, Hoban S, Kelley JL, Lotterhos KE, Reed LK, Antolin MF, Storfer A. 2017. Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Mol Ecol Resour*. 17:142–152.

Lowry DB, Willis JH. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol*. 8:e1000500.

Macqueen DJ, Johnston IA. 2014. A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proc Biol Sci*. 281:20132881.

Marques DA, Lucek K, Meier JI, Mwaiko S, Wagner CE, Excoffier L, Seehausen O. 2016. Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genet*. 12:e1005887.

Martínez A, Garza JC, Pearse DE. 2011. A microsatellite genome screen identifies chromosomal regions under differential selection in steelhead and rainbow trout. *Trans Am Fish Soc*. 140:829–842.

Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. 2008. Database indexing for production MegaBLAST searches. *Bioinformatics*. 24:1757–1764.

Natarajan C, Hoffmann FG, Weber RE, Fago A, Witt CC, Storz JF. 2016. Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science*. 354:336–339.

Nichols KM, Broman KW, Sundin K, Young JM, Wheeler PA, Thorgaard GH. 2007. Quantitative trait loci x maternal cytoplasmic environment interaction for development rate in *Oncorhynchus mykiss*. *Genetics*. 175:335–347.

Nichols KM, Edo AF, Wheeler PA, Thorgaard GH. 2008. The genetic basis of smoltification-related traits in *Oncorhynchus mykiss*. *Genetics*. 179:1559–1575.

Noor MA, Bennett SM. 2009. Islands of speciation or mirages in the desert? Examining the role of restricted recombination in maintaining species. *Heredity (Edinb)*. 103:439–444.

Northcote TG. 2010. Controls for trout and char migratory/resident behaviour mainly in stream systems above and below waterfalls/barriers: a multidecadal and broad geographical review. *Ecol Freshw Fish*. 19:487–509.

Nosil P, Feder JL. 2012. Genomic divergence during speciation: causes and consequences. *Philos Trans R Soc Lond B Biol Sci*. 367:332–342.

Ohno S. 1970. *Evolution by gene duplication*. New York: Springer-Verlag.

Orr HA. 2005. The genetic theory of adaptation: a brief history. *Nat Rev Genet*. 6:119–127.

Pauling L, Itano HA. 1949. Sickle cell anemia a molecular disease. *Science*. 110:543–548.

Pearse DE. 2016. Saving the spandrels? Adaptive genomic variation in conservation and fisheries management. *J Fish Biol*. 89:2697–2716.

Pearse DE, Barson NJ, Nome T, Gao G, Campbell MA, Abadía-Cardoso A, Anderson EC, Rundio DE, Williams TH, Naish KA, *et al.* 2019. Sex-dependent dominance maintains migration supergene in rainbow trout. *Nat Ecol Evol*. 3:1731–1742.

Pearse DE, Campbell MA. 2018. Ancestry and adaptation of rainbow trout in Yosemite National Park. *Fisheries*. 43:472–484.

Pearse DE, Hayes SA, Bond MH, Hanson CV, Anderson EC, Macfarlane RB, Garza JC. 2009. Over the falls? Rapid evolution of ecotypic differentiation in steelhead/rainbow trout (*Oncorhynchus mykiss*). *J Hered*. 100:515–525.

Pearse DE, Miller MR, Abadía-Cardoso A, Garza JC. 2014. Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proc Biol Sci*. 281:20140012.

Perrier C, Bourret V, Kent MP, Bernatchez L. 2013. Parallel and nonparallel genome-wide divergence among replicate population pairs of freshwater and anadromous Atlantic salmon. *Mol Ecol*. 22:5577–5593.

Phillis CC, Moore JW, Buoro M, Hayes SA, Garza JC, Pearse DE. 2016. Shifting thresholds: rapid evolution of migratory life histories in steelhead/rainbow trout, *Oncorhynchus mykiss*. *J Hered*. 107:51–60.

Poh YP, Domingues VS, Hoekstra HE, Jensen JD. 2014. On the prospect of identifying adaptive loci in recently bottlenecked populations. *PLoS One*. 9:e110579.

Prince DJ, O'Rourke SM, Thompson TQ, Ali OA, Lyman HS, Saglam IK, Hotaling TJ, Spidle AP, Miller MR. 2017. The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. *Sci Adv*. 3:e1603198.

Quinn TP. 2005. *The behavior and ecology of Pacific salmon and trout*. Seattle, WA: University of Washington Press.

Rastogi S, Liberles DA. 2005. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol Biol*. 5:28.

Ravinet M, Westram A, Johannesson K, Butlin R, André C, Panova M. 2016. Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Mol Ecol*. 25:287–305.

Reid NM, Proestou DA, Clark BW, Warren WC, Colbourne JK, Shaw JR, Karchner SI, Hahn ME, Nacci D, Oleksiak MF, *et al.* 2016. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*. 354:1305–1308.

Roesti M, Gavrilets S, Hendry AP, Salzburger W, Berner D. 2014. The genomic signature of parallel adaptation from shared genetic variation. *Mol Ecol*. 23:3944–3956.

Ruegg K, Anderson EC, Boone J, Pouls J, Smith TB. 2014. A role for migration-linked genes and genomic islands in divergence of a songbird. *Mol Ecol*. 23:4757–4769.

Rundio DE, Garza JC, Lindley ST, Williams TH, Pearse DE. 2020. Differences in growth and condition of juvenile *Oncorhynchus mykiss* related to sex and a migration-associated genomic region. *Can J Fish Aquat Sci*. 78:322–331.

Salisbury SJ, McCracken GR, Perry R, Keefe D, Layton KKS, Kess T, Nugent CM, Leong JS, Bradbury IR, Koop BF, *et al.* 2020. Limited genetic parallelism underlies recent, repeated incipient speciation in geographically proximate populations of an Arctic fish (*Salvelinus alpinus*). *Mol Ecol*. 29:4280–4294.

Schwander T, Libbrecht R, Keller L. 2014. Supergenes and complex phenotypes. *Curr Biol*. 24:R288–R294.

Sémon M, Wolfe KH. 2008. Preferential subfunctionalization of slow-evolving genes after allopolyploidization in *Xenopus laevis*. *Proc Natl Acad Sci U S A*. 105:8333–8338.

Sheer MB, Steel EA. 2006. Lost watersheds: barriers, aquatic habitat connectivity, and salmon persistence in the Willamette and Lower Columbia River Basins. *Trans Am Fish Soc*. 135:1654–1669.

Sundin K, Brown KH, Drew RE, Nichols KM, Wheeler PA, Thorgaard GH. 2005. Genetic analysis of a development rate QTL in backcrosses of clonal rainbow trout, *Oncorhynchus mykiss*. *Aquaculture*. 247:75–83.

Thompson NF, Anderson EC, Clemento AJ, Campbell MA, Pearse DE, Hearsey JW, Kinziger AP, Garza JC. 2020. A complex phenotype in salmon controlled by a simple change in migratory timing. *Science*. 370:609–613.

Tuttle EM, Bergland AO, Korody ML, Brewer MS, Newhouse DJ, Minx P, Stager M, Betuel A, Cheviron ZA, Warren WC, *et al.* 2016. Divergence and functional degradation of a sex chromosome-like supergene. *Curr Biol*. 26:344–350.

Van de Peer Y, Mizrachi E, Marchal K. 2017. The evolutionary significance of polyploidy. *Nat Rev Genet*. 18:411–424.

Warren IA, Ciborowski KL, Casadei E, Hazlerigg DG, Martin S, Jordan WC, Sumner S. 2014. Extensive local gene duplication and functional divergence among paralogs in Atlantic salmon. *Genome Biol Evol*. 6:1790–1805.

Wellenreuther, M., Bernatchez, L. 2018. Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecol. Evol.* 2379:1–14.

Winans GA, Allen MB, Baker J, Lesko E, Shrier F, Strobel B, Myers J. 2018. Dam trout: genetic variability in *Oncorhynchus mykiss* above and below barriers in three Columbia River systems prior to restoring migrational access. *PLoS One*. 13:e0197571.

Wolf JB, Ellegren H. 2017. Making sense of genomic islands of differentiation in light of speciation. *Nat Rev Genet*. 18:87–100.