**Genetic diversity within late-summer run and half-pounder steelhead (*Oncorhynchus mykiss)* in the Rogue River, Oregon**

David I Dayan1\*, Steve Mazur3, Laura J Green3, Andrew J Wells3, Marc A Johnson2, Daniel J Van Dyke4, Peter A Samarin4, Ryan D Battleson4, Kathleen G O’Malley1

1 State Fisheries Genomics Laboratory, Coastal Oregon Marine Experiment Station, Department of Fisheries, Wildlife and Conservation Sciences, Hatfield Marine Science Center, Oregon State University, Newport, OR, USA

2 Native Fish Conservation and Recovery, Oregon Department of Fish and Wildlife, Salem, OR, USA

3 Oregon Department of Fish and Wildlife, Gold Beach, OR, USA

4 Rogue Watershed District, Oregon Department of Fish and Wildlife, Central Point, OR, USA

\* Corresponding author, email: [david.dayan@oregonstate.edu](mailto:david.dayan@oregonstate.edu), ORCID ID: 0000-0001-5001-4062

**Keywords**

Migration timing, life history, anadromy, Pacific salmon, *Oncorhynchus,* population genetics, *greb1l - rock1*

**Acknowledgments**

We thank the Oregon State University Center for Quantitative Life Sciences and the University of Oregon Genomics & Cell Characterization Core Facility for DNA sequencing the Rogue River steelhead samples and for assisting with their computational infrastructure. We thank Sandra Bohn for her contribution to data collection. We thank Peter Stevens and Chris Lorion from the Oregon Department of Fish and Wildlife for their helpful comments on the manuscript.

**Abstract**

Anadromous *Oncorhynchus mykiss* (steelhead) express diverse migratory behaviors and life history strategies. On the Rogue River in Oregon, adult steelhead return from the ocean during distinct early-summer, late-summer, and winter runs. Additionally, some juvenile steelhead briefly return to freshwater as “half pounders,” before completing their oceanic growth phase and returning to freshwater again to spawn. Using markers from a Genotyping‐in‐Thousands by sequencing amplicon panel, we describe a genetic axis that discriminates between Rogue River steelhead that express early- *vs.* late- migration phenotypes. Then we examine whether late-summer run adults carry early- or late-migration associated alleles and assign half pounders as early or late migrators. Both late-summer run adults and half pounders bear highly heterozygous genotypes and recombined haplotypes at the *greb1l-rock1* region on chromosome 28 associated with adult migration timing, suggestive of ongoing gene flow among individuals with early- and late-migration haplotypes. Our classification suggests that half pounders are a mixed assemblage of offspring from all three adult runs. We discuss the utility of markers within the *greb1l-rock1* region to predict phenotypes and highlight the need to validate candidate gene-trait associations across lineages before applying them for management. Finally, we consider the implications of our results on the maintenance of life history diversity within steelhead.

**Introduction**

Maintaining phenotypic and genetic diversity within species is a primary goal of conservation and management (Mimura et al. 2017). Variation in ecologically relevant traits provides short-term resilience against environmental changes and may promote long-term survival through evolutionary adaptation (Merilä and Hendry 2014; Schindler et al. 2015). Pacific salmonids (*Oncorhynchus* spp.) exhibit tremendous life history diversity with anadromous forms undertaking oceanic feeding migrations prior to returning to their natal freshwater environment to spawn (Groot and Margolis 1991). The timing of adult migration to freshwater varies within species and has been used for more than a century to define salmon and steelhead stocks (Kostow 2009). This portfolio of adult migration timing variation is a major focus of researchers and managers hoping to take advantage of the relationship between diversity and stability (Greene et al. 2010; Schindler et al. 2010; Satterthwaite and Carlson 2015; Ford et al. 2020).

While migratory phenotypes in Pacific salmonids are complex, variation within some rivers can be characterized by two adult migratory types. Spring-returning Chinook salmon (*O. tshawytscha*) and summer-returning steelhead (*O. mykiss*) are referred to as “early migrators,” while Chinook salmon that return in fall and steelhead that return in winter are considered “late migrators.” Early migrators tend to begin their migrations in a sexually immature state, and mature while holding in freshwater habitats, while late migrators arrive in freshwater with more developed gonads and rapidly ascend rivers shortly before spawning (Hearsey and Kinziger 2015). Early migration is associated with numerous behavioral, physiological, and morphological traits that may facilitate holding through summer months and permit access to distinct freshwater habitats farther from the ocean and at higher elevations. However, early migration imposes substantial costs (Quinn et al. 2016). Individuals must arrive with enough stored energy to survive until spawning in upstream habitats that expose them to additional risk of disease, predation, and unfavorable environmental conditions. While complete allopatry is uncommon, upstream spawning habitats preferentially used by early migrators are also particularly impacted by human activities through changes to passage barriers and stronger impacts of climate change (Beechie et al. 2006; McClure et al. 2008; Sheer and Steel 2011; Crozier et al. 2019).

In addition to promoting stability and long-term viability by contributing to the portfolio of variation, early-migrating Chinook salmon and steelhead are ecologically, culturally and socially significant. Early migrators bring marine-derived nutrients higher into watersheds and are available as prey for longer than late migrators (Quinn et al. 2016), reshaping the ecosystems they use for spawning (Warren and McClure 2012). The high fat content of early migrators makes them a prized catch for anglers. Importantly, maintaining runs of early migrators enhances food security for indigenous people by extending fishing seasons and contributing to fishery stability (Swezey 1977; Nesbitt et al. 2016). Therefore protecting migration timing diversity is a critical component in the movement towards indigenous food sovereignty (Jacob et al. 2010; Coté 2016).

Some Pacific salmonids express variation in migratory behavior beyond these generalized early- and late-migrating adult types. Freshwater entry timing can be quite broad in major rivers, spanning more than two distinct pulses of individuals. Also, unlike most other Pacific salmonids, steelhead are iteroparous (i.e. undergo repeat spawning), and repeat spawning rates vary among populations (Keefer et al. 2008). Some steelhead migrate to the ocean, but only spend 2-5 months at sea before returning to freshwater. Most of these “half pounders” are sexually immature and therefore will migrate back to the ocean and return to freshwater later as mature adults to spawn (Kesner and Barnhart 1972). The half-pounder life history is observed in northern California, USA (Klamath, Eel and Mad Rivers)(Kesner and Barnhart 1972), southern Oregon, USA (Rogue River)(Everest 1973), the Kamchatka Peninsula, Russia (Utkholok River)(McPhee et al. 2007), and in some introduced populations in Argentina (Pascual et al. 2001). Relative to most returning adult steelhead, half pounders exhibit small size, short upstream migrations, sexual immaturity, and active freshwater feeding (Kesner and Barnhart 1972; Everest 1973; Hodge et al. 2014).

Adult steelhead from the Rogue River, Oregon express substantial diversity in freshwater entry timing. Early observations suggest that historical freshwater entry time may have spanned all months of the year, and upstream passage of Gold Ray Dam at river kilometer 202 was recorded in all months of most years (Rivers 1963). Gaps between three distinct peaks of freshwater entry timing became apparent after population declines in the early twentieth century and led to the recognition of early, intermediate and late runs, initially termed spring, fall and winter, respectively (Rivers 1963). The extended period of freshwater entry and the observation of three, not two, distinguishable peaks of freshwater entry timing in the Rogue River prompted a longstanding question: are Rogue steelhead with intermediate migration timing better understood as summer steelhead (i.e. early migrators), winter steelhead (i.e. late migrators) or a distinct group? Similar questions were posed throughout the broader region, including the Klamath River, where steelhead with both intermediate migration timing and half pounders are found (Roelofs 1983; Busby 1994, 1996).

While there is overlap in the timing and habitats used for spawning between the early and intermediate runs of Rogue River steelhead, early run individuals tend to arrive on spawning grounds earlier and preferentially use spawning grounds further upstream than intermediate run individuals (Rivers 1963; Everest 1971, 1973). Mark-recapture experiments conducted on iteroparous individuals revealed that early and intermediate run steelhead had a tendency to express the same migration timing phenotype during their second spawning run (Everest 1973). Hatchery experiments demonstrated that early and intermediate run steelhead tend to produce offspring of their own run-type (Everest 1971, 1973), although others have cast doubt on this finding (Leider 1985). Ultimately, observations of tagged early and intermediate steelhead together on spawning grounds led fisheries biologists to conclude that there was insufficient genetic isolation to warrant management of the early and intermediate runs as separate groups (Everest 1971, 1973). Both early and intermediate runs were thereafter treated as a single summer run (Everest 1971, 1973; Roelofs 1983). Terminology for the early and intermediate runs on the Rogue River changed to reflect this classification. Steelhead entering the river during the first peak became known as the early-summer run and those entering in the second peak as the late-summer run (Everest 1971, 1973; Roelofs 1983). This view was subsequently supported by the absence of genetic differentiation at allozymes (Reisenbichler et al. 1992).

Winter steelhead return to the Rogue River in the third peak. Winter steelhead spawn later in the season and in more downstream portions of the river than early- and late-summer steelhead, although this separation is also not complete. The extent of gene flow owing to partial spatiotemporal overlap in spawning among summer and winter steelhead on the Rogue River remains unclear (ODFW 1990, 1994).

Freshwater entry timing of half-pounders in the Rogue River typically overlaps strongly with late-summer run adults and peaks in late August (Everest 1971, 1973; ODFW 1994). A substantial portion of adult steelhead in both the summer and winter runs express the half-pounder life history (Everest 1973; ODFW 1990; Evenson and Ewing 1992; ODFW 1994). While nearly all half pounders are immature, varying proportions of both precocious males and mature females have been observed in the Rogue River (Everest 1973), and up to 8% are expected to spawn during their half-pounder run in the Klamath River, CA (Hodge et al. 2014). Neither the proximate genetic and physiological mechanisms, nor the ultimate evolutionary forces underlying the maintenance of the half-pounder phenotype are well understood (Cramer 1985; Busby 1994; Hopelain 1998; Hodge et al. 2014; Peterson et al. 2017).

Recent genomic studies have led to major advances in our understanding of the genetic architecture of adult migration timing and associated phenotypes in Chinook salmon and steelhead. While most ecologically relevant complex phenotypes are expected to have a polygenic basis (Rockman 2012; Boyle et al. 2017), association studies across multiple *Oncorhynchus* lineages have pointed to a single genomic region on chromosome 28, near the *greb1l* and *rock1* genes, responsible for the majority of variation in adult migration timing (Hess et al. 2016)(reviewed in Waples et al. 2022). The apparent oligogenic architecture underpinning adult migration timing has amplified conservation concerns (Prince et al. 2017) and prompted a number of important questions for future research (Waples and Lindley 2018; Oke and Hendry 2019; Ford et al. 2020; Waples et al. 2022). While the majority of work to date has focused on the extremes along the continuum of early and late migration timing, we believe investigations into the diversity of migratory phenotypes, including the intermediate late-summer steelhead, as well as the half-pounder life history, could provide meaningful insights relevant to the conservation of salmonid life history diversity (Pearse 2019; Ford et al. 2020; Thompson et al. 2020).

In this study, we investigate genetic variation within late-summer run and half-pounder steelhead in the Rogue River. We ask if late-summer adults bear early- or late-migration associated alleles or haplotypes, and if half pounders can be genetically classified as early or late migrators. Our study also adds to the growing body of research cataloguing haplotype diversity in the *greb1l-rock1* region and examines the extent to which individuals with intermediate migratory phenotypes possess heterozygous genotypes and can serve as a reservoir of early migration associated alleles.

**Methods**

Reproducible Research

Detailed logs and all data needed to replicate this study are available at a github repository: <https://github.com/david-dayan/half_pounder_2023> archived at zenodo with a stable identifier ([DOI:10.5281/zenodo.7931542](https://zenodo.org/badge/latestdoi/636851663)). R notebooks containing narrative logs of all analyses with integrated code, results and commentary are available as supplementary files, including generating filtered datasets from raw sequencing reads (Online Resource 1) and conducting analyses (Online Resource 2).

Sampling and Migration Timing Groups

Throughout this study we consider four phenotypic groups of Rogue River steelhead, including the juvenile half pounders and three adult groups. We categorize adult steelhead into early-summer, late-summer or winter runs based on freshwater entry timing (Table 1). These categories represent distinguishable phenotypic groups because freshwater entry timing can be clustered into three peaks, but do not necessarily represent distinct genetic lineages.

Oregon Department of Fish and Wildlife (ODFW) staff collected caudal fin clips from Rogue River steelhead from 2018-2020. Early-summer steelhead on the Rogue River are defined as those that enter during the first peak of freshwater entry that occurs sometime during May, June, or July. Tissue samples from early-summer run individuals were collected at the Cole M. Rivers Hatchery sorting pond (river kilometer 253) on June 26th, 2019, indicating their freshwater entry timing was substantially before the subsequent late-summer peak that occurs sometime during August or October (Table 2). Tissue samples from late-summer run steelhead were collected during ODFW’s Huntley Park Seining Project just above the estuary at river kilometer 13 (Huntley Park, Fig. 1) in 2019 and 2020. Seining at Huntley Park began on July 15th and the first individuals from the Huntley Seining Project in 2019 and 2020 were collected on July 26th and July 29th, respectively. The final individuals included in the late-summer run collection in 2019 and 2020 were collected on October 21st and October 16th, respectively. We included the five individuals sampled in late July alongside the 270 individuals sampled in August, September, or October as the late-summer run collection (Table 2), although we note that these five individuals were sampled earlier than the late-summer peak of freshwater entry in these years. Winter run steelhead enter the Rogue River from November through May. Tissue samples from winter run steelhead were collected at the Cole M. Rivers Hatchery sorting pond and from the Applegate Dam trap (river kilometer 75 of the Applegate River, a tributary of the Rogue at river kilometer 154) in 2019 on May 1st and May 13th (Table 2).

Tissue samples from half pounders were also collected during the Huntley Seining Project (Fig. 1) in 2018 and 2019. Half pounders were identified as individuals with fork lengths between 250 - 410mm. Sample sizes, sex, and origin (hatchery-origin (HOR) *vs.* natural origin (NOR)) is presented in Table 2 and in a supplemental file (Online Resource 3).

Genotyping

We genotyped individuals at a panel of previously described single nucleotide polymorphisms (SNPs) using Genotyping-in-Thousands by sequencing (GT-seq) (Campbell et al. 2015). We began filtering genotypes based on missingness, and the individual fuzziness index (IFI), which estimates the amount of cross-contamination in a given sample. We took an iterative approach to missingness and IFI filtering and recalculated missingness for all individuals and genotypes between each step. In the final step, we removed individuals with >10% missing data, markers with >20% missing data, and individuals with IFI >2.5. We then removed sites with poorly calibrated allele correction values or more than three clusters of allele ratios suggestive of a paralogous sequence variant, and any monomorphic markers. We also removed any markers that displayed strong group-specific patterns of missingness. Finally, we excluded likely duplicate tissue samples that arise due to batch sampling using Coancestry (Wang 2011). Complete details of genotyping methods are available in the genotyping notebook (Online Resource 1).

GT-seq Panel and Genetic Datasets

The GT-seq panel used in this study was developed by the Columbia River Inter-Tribal Fish Commission (Collins et al. 2020) and includes 391 markers. The panel contains 251 neutral markers, 136 putatively adaptive markers, three species identification markers, and a sex identification marker. Putatively adaptive markers were identified from environmental association analyses, or as markers that demonstrate strong associations with ecologically relevant phenotypic variation (Pearse et al. 2014; Hess et al. 2016; Micheletti et al. 2018a; Micheletti et al. 2018b). Marker-level information regarding the GT-seq panel is included as a supplemental file (Online Resource 4).

We conducted analyses on two datasets: neutral and migration timing. To generate the “neutral dataset”, we excluded one marker from any pair that map to the same chromosome or contig with an r2 > 0.2. The migration timing dataset was composed of genotypes that map to chromosome 28 of the *O. mykiss* genome in the *greb1l-*rock1 and passed *.* This region is strongly associated with adult migration timing across multiple populations (Hess et al. 2016; Prince et al. 2017; Micheletti et al. 2018a; Collins et al. 2020). Markers that passed all genotype quality filtering steps and were included in either the neutral or migration timing dataset are noted in Online Resource 4.

Genetic Diversity

We calculated per-marker and average observed and expected heterozygosity at both the neutral and migration timing datasets using *adegenet* (Jombart and Ahmed 2011). Hardy-Weinberg proportion significance testing at individual markers was conducted using *pegas* (Paradis 2010). We used ﻿an exact test based on 1000 Monte Carlo permutations, and corrected the resulting p-values for multiple comparisons using the FDR.

Genetic Structure and Classification

We ran STRUCTURE (Falush et al. 2003) on the neutral dataset with an admixture model, correlated allele frequency, a burn-in of 10,000 iterations, followed by 20,000 iterations and conducted 10 replicates for one to five putative ancestral genetic clusters (*k*). We estimated best *k* using the ∆K method of Evanno (Evanno et al. 2005). However, given the potentially low level of differentiation in our dataset and uneven sampling, we considered STRUCTURE output across multiple *k* as potentially biologically meaningful and included these results (Cullingham et al. 2020; Stankiewicz et al. 2021). Replicate results within each *k* were combined using the clumpak algorithm (Kopelman et al. 2015) on the clumpak webserver. We conducted principal component analysis (PCA) on the neutral dataset using the *ade4* package (Thioulouse and Dray 2007). Missing data were imputed using mean allele frequency prior to ordination. We also grouped individuals by phenotype (late-summer, half-pounder, early-summer and winter run), then estimated pairwise differentiation (FST) between groups using Weir and Cockerham’s estimator (1984), and calculated other F-statistics at both the neutral and migration timing datasets using the *hierfstat* package.

We used *a priori* DAPC to classify late-summer run adults and half pounders as genetically similar to early migrators, late migrators or unassigned. We trained the DAPC using only early-summer and winter individuals to identify a linear combination of alleles in the migration timing dataset that strongly discriminated between these groups (LD1). Then we assigned late-summer and half-pounder individuals into bins of early migrators, late migrators or unassigned on the basis of their individual scores along LD1. The number of retained principal components was chosen using cross validation with 100 replicates at each PC and 9:1 ratio of training to test datasets. Bins for assignment were bounded by the maximum and minimum observed values among the early-summer and winter run collections. To understand if there was a temporal cline in migration timing alleles during the late-summer run, we fit a linear model of Julian date of sampling at river kilometer 13 to approximate freshwater entry timing. We used individual score on LD1 as a fixed effect and sample year as a covariate. We fit the model*,* examined residual plots for deviations from homoscedasticity or normality, and evaluated significance of the effect of individual score on sampling date after controlling for year using an F-test.

Migration Timing Haplotypes

To examine haplotype diversity and structure among our collections at the *greb1l-rock1* region, we examined patterns of linkage disequilibrium between markers and statistically phased genotypes using *fastPHASE* v1.4.8 (Scheet and Stephens 2006). We ran *fastPHASE* using default settings on genotypes at the eight migration timing markers in the migration timing dataset that showed strong differences in allele frequency between early-summer and winter runs. To visualize the relationships among these inferred haplotypes, we constructed a haplotype network using a minimum spanning network in Popart v1.7 (Leigh and Bryant 2015), and hierarchically clustered haplotypes within each of the four phenotypic groups.

**Results**

Genotypes

After genotype quality control, the dataset consisted of 1000 individuals genotyped at 350 markers. Median read depth per individual per locus after filtering was 191, and the mean was 328 ± 428 (s.d.). Further detail of filtering results are available in the genotyping notebook (Online Resource 1). There were 237 neutral markers after genotype quality control. After LD-thinning, there were 233 markers in the neutral dataset (Online Resource 4). There were 12 markers in the migration timing dataset (Online Resource 4). All groups contained more females than males based on the genetic sex marker (Table 2), but half pounders had the least unbalanced sex ratio.

Genetic Diversity and Differentiation

Mean expected heterozygosity (He) in the neutral dataset ranged from 0.304 – 0.311 across runs (Supplemental Table 1).

At the migration timing markers, He was higher in late-summer run (He = 0.355) and half pounders (He = 0.357) than in early-summer (He = 0.009) and winter runs (He = 0.096) (Supplemental Table 1). No markers demonstrated significant departures from Hardy-Weinberg proportions within early-summer or winter collections. Migration timing markers were within Hardy-Weinberg proportion in all adult collections, but there was excess homozygosity at 9 of the 12 markers in the half-pounder collection (FDR-adjusted p-values of Monte-Carlo simulation < 0.05, 1000 permutations per marker per population).

There was relatively limited differentiation between half pounders sampled in 2018 and half pounders sampled in 2019 using either the neutral or migration timing dataset (FST = 0.0006 and FST = 0.0150, respectively). Interannual differentiation between collections of late-summer run was also low (FST = 0.0007 and FST = 0.0000, respectively). Therefore, we pooled our collections across years in all other analyses. Overall, FST using the neutral dataset was small (FST = 0.0047). Pairwise FST in the neutral dataset ranged from 0.0002 – ﻿0.0105 (Table 3) and was strongly correlated with pairwise FST in the migration timing dataset (Mantel test statistic = 0.85, p-value = 0.04, Mantel test with 999 permutations). Pairwise FST in the migration timing dataset ranged from near zero between the late-summer and half-pounders collections, to near fixation (FST = 0.925) between early-summer and winter runs (Table 3).

Genetic Structure

*PCA*

PCA using the neutral dataset did not reveal distinct clusters of genetic variation within our collections of Rogue River steelhead (Fig. 2). Examination of the PCA screeplot suggested only the first principal component (PC) should be considered. Late-summer run and half pounders were not distinguishable from one another along the first PC; both the centroids and extents of the 95% confidence intervals along the first PC for these two phenotypic groups were very similar (Fig. 2). The first PC described genetic variation within late-summer run and half pounders that was mostly absent from either early-summer or winter runs.

*STRUCTURE*

The most likely number of ancestral genetic clusters within our collections, according to the Evanno method (Evanno et al. 2005), was two (*k* = 2). STRUCTURE results using the neutral dataset modeling *k* of three or more demonstrated similar patterns (Fig. 3, Supplemental Fig. 1). Regardless of *k* value, there was considerable variation at the individual level suggestive of a high degree of admixture, or limited differentiation between clusters. When using three or more ancestral genetic clusters, there was a cluster that contributed little to the ancestry of early-summer and winter run collections, but substantially to late-summer run and half-pounder collections. This cluster was also more variable within late-summer and half-pounders collections than in the early-summer and winter run collections. A summary of the mean and variation of ancestry proportions for *k =* 3 and *k =* 4 is available in Supplemental Tables 2 and 3.

Classification and Clinal Analysis

We assigned individuals from our collection of late-summer run adults and half pounders as genetically similar to early migrators, late migrators or unassigned. A single principal component was retained in the DAPC. Therefore, LD1 was equivalent to the first principal component of a PCA of early-summer and winter individuals. For late-summer run individuals, the majority were unassigned (78%), many assigned as late migrators (21%) and a small number assigned as early migrators (1%) (Table 4, Fig. 4). For half-pounder individuals, approximately one half were unassigned (53%), many assigned as late migrators (35%) and a small number assigned as early migrators (12%) (Table 4, Fig. 4).

There was a temporal pattern in allele frequency from more early-migration to late-migration associated alleles during the late-summer run (Fig. 5), suggestive of clinal variation in migration timing. Individual score on LD1 significantly explained variation in sampling date, after controlling for differences in sampling date between years (p-value = 0.0001, F-test), but the estimated effect size was small: 0.197 days per unit of LD1. This effect size was equivalent to 12.9 days from maximum observed early-migration score to maximum observed late-migration score among the late-summer individuals.

Migration Timing Markers

We investigated allele frequency variation and haplotype structure at the *greb1l-rock1* region of chromosome 28 that is associated with adult migration timing and densely genotyped by our GT-seq panel. Of the 12 migration timing markers included in our study, four proved to be uninformative of migration timing: the major allele was the same in both early-summer and winter collections (Fig. 6), and allele frequency did not vary between early-summer and winter collections. Seven of the 12 migration timing markers were fixed for alternative alleles in the early-summer and winter run collections. An eighth marker was highly informative, but not fixed in the winter run collection (Fig. 6). These eight markers loaded heavily on LD1, while the remaining four had near zero loading. Our collections of late-summer run adults and half pounders demonstrated intermediate allele frequencies at these eight informative markers. However, in addition to individuals with consistently early-migration associated homozygous genotypes, late-migration associated homozygous genotypes, or heterozygous genotypes across these markers, many individuals displayed recombinant patterns that were homozygous for alternatively early- and late-migration associated alleles at adjacent positions along the genome. To examine this apparent disruption of the strong linkage observed in early-summer and winter runs in the *greb1l-rock1* region (Supplemental Fig. 2), we statistically phased genotypes at these eight markers.

We inferred 22 haplotypes in the late-summer run, 25 in half pounders, a single haplotype in the early-summer run, and two in winter run. To better understand the relationships among these haplotypes we inferred a haplotype network (Supplemental Fig. 3). There were no private haplotypes within either the early-summer or winter run collections; many late-summer run and half pounders carry complete early- or late-migration haplotypes with no recombination. To examine how haplotype diversity was distributed among groups of late-summer run adults and half-pounders assigned as early- or late-migrators or unassigned, we hierarchically clustered haplotypes within each of these groups. Late-summer run adults and half-pounders that assigned as genetically similar to early- or late-migrators carried two copies each of the complete early- or late-migration haplotypes, respectively. Among unassigned late-summer run adults and half-pounders, some individuals carried a single copy of either the complete early- and late-migration haplotypes, but recombination was common (Fig. 7).

**Discussion**

Structure at Neutral Markers

We found no evidence of strong genetic structure at neutral markers. Our PCA and STRUCTURE results both suggest that there is greater variation at a subset of neutral markers within our collections of late-summer run and half pounders than early-summer and winter runs. We attribute this pattern to our sampling design; the minor genetic structure observed is likely due to restricted gene flow among spawning groups from different tributaries upstream of the estuary where late-summer and half-pounder individuals were sampled. Our conclusion is also supported by the observation that multiple neutral markers demonstrated an excess of homozygosity within our late-summer run and half-pounder collections, but not our early-summer or winter run collections, suggestive of Wahlund effects in the former groups. We might have observed a similar pattern in early-summer and winter runs had they also been sampled closer to the estuary, or if upstream sampling for early-summer and winter runs had been conducted across more tributaries.

Are Late-Summer Run Steelhead better Understood as Genetically Early or Late Migrators?

On the Rogue River, adult steelhead enter freshwater in three distinct pulses with relatively fewer entering freshwater entry in the periods between them. Equipped with new genetic tools, we revisit a longstanding question (Everest 1973; Roelofs 1983; Busby 1994): Are Rogue River steelhead with intermediate migration timing better understood as summer steelhead (i.e. early migrators), winter steelhead (i.e. late migrators) or something else? Our results paint a more complex picture than common understanding, highlight the challenges associated with clustering biological variation into discrete categories, and have implications for the conservation of life history diversity within steelhead.

While we assigned some late-summer run individuals as genetically similar to early or late migrators, most late-summer run individuals went unassigned. Extended sampling within the early-summer and winter runs may reveal additional variation at migration timing markers within these groups and widen the intervals used to assign late-summer individuals into early- or late-migrator clusters. However, the distribution of late-summer run steelhead along the early- *vs.* late-migration genetic axis was not bimodal. Most late-summer run fish demonstrated intermediate scores along this axis, due to high heterozygosity at migration timing markers or recombination between migration timing haplotypes. Indeed, direct examination of inferred migration timing haplotypes confirmed that many unassigned late-summer run individuals possess a single copy each of early- and late-migration haplotypes, suggesting that some unassigned fish are first generation hybrids between parents homozygous for either the early- or late-migration associated haplotypes. However, recombined haplotypes among the unassigned late-summer run individuals were also common. Recombination between migration timing haplotypes implies gene flow between individuals that carry early- and late-migration haplotypes.

Taken together, our results suggest that late-summer run Rogue River steelhead as a whole cannot be genetically classified as either early- or late-migrators. Instead, ongoing gene flow and resulting recombination produces a continuous range of genetic variation within late-summer run steelhead. Although a more thorough understanding of the genotype to phenotype map would be required to accurately predict effects from recombined migration timing haplotypes, our results suggest that genetic variation within the *greb1l-rock1* region can produce a temporal cline in freshwater entry timing in steelhead. Consequently, there may be no strict genetic boundaries between runs. However, the observation of a genetic migration timing cline must be placed within the context of the ecology of steelhead on the Rogue River. The freshwater entry timing phenotype is not continuous, but clustered into three distinct runs. Regardless of whether clustering of freshwater entry timing into peaks is due to exogenous environmental or endogenous genetic factors, this discontinuity will always preclude clear parsing of distinct genetic migration timing clusters from a continuous early- to late-migration genetic cline.

Half Pounders

We found little evidence of genetic differentiation between our collections of adult late-summer run and the predominately immature half-pounder steelhead that enter the river at the same time of year. Therefore, much of our discussion of structure within late-summer runs, above, is also applicable to half pounders*.* However, we emphasize that late-summer run adults are returning to spawn, while most half pounders will not spawn until they return to freshwater as adults in subsequent years. Also, we found excess homozygosity at migration timing markers within our collection of half pounders, but not among late-summer run adults. The most parsimonious interpretation of our results therefore differs between late-summer run and half pounders; half pounders are a mixed assemblage of individuals that are likely the offspring of early-summer, late-summer and winter run adults.

Conservation Implications

The recent finding of oligogenic architecture of adult migration timing has amplified existing conservation concerns around early-migrating salmon and steelhead and motivated many research questions (Kardos and Shafer 2018; Waples and Lindley 2018; Oke and Hendry 2019; Ford et al. 2020; Waples et al. 2022). We attempt to address some of these with our data. First, from a management perspective, identification of genetic markers that can be used to reliably predict the expression of early- and late-migrating phenotypes is a priority. Allele frequency varied strongly between our relatively small early-summer *vs.* winter run Rogue River collections at only 8 of the 12 genetic markers with known migration timing associations in steelhead from the Columbia River (Hess et al. 2016; Micheletti et al. 2018a; Micheletti et al. 2018b). This result is consistent with previous studies that have identified varying strengths of association with adult migration timing phenotypes (Willis et al. 2020) and lineage-specific patterns of linkage disequilibrium within the *greb1l-rock1* region in steelhead (Collins et al. 2020). Our findings highlight the importance of lineage-specific marker validation prior to use in management.

Additional gaps in understanding include whether reservoirs of early migration associated alleles are available among late-migrating populations, patterns of dominance, and the direction and magnitude of gene flow between early- and late-migrating runs (Waples and Lindley 2018; Ford et al. 2020; Oomen et al. 2020). Our focus on individuals with intermediate phenotypes permits important insights into these questions. The late-summer run in the Rogue River is primarily composed of heterozygotes and fish with recombined migration timing haplotypes. This suggests that there is not strong dominance of migration timing alleles in this lineage; individuals with intermediate genotypes tend to express intermediate phenotypes. Spatial differences in spawning habitat between early- and late-summer runs on the Rogue may produce differences in the environmental conditions experienced by embryos and juveniles (Everest 1971, 1973), and late-summer adults spend less time exposed to freshwater stressors. Together, these observations point to late-summer run fish as particularly important for conservation on the Rogue River, because they may serve as a temporary reservoir of early-migration associated alleles that are buffered from the stronger selection imposed on the early-summer run (Thompson et al. 2019). Currently, fishery managers recognize late-summer Rogue River steelhead as summer steelhead, which are considered a sensitive species in Oregon, warranting focused protections. Nearby steelhead populations in the Eel and Klamath Rivers also exhibit considerable heterozygosity at the *greb1l-rock1* region and migration timing diversity (Pearse 2019; Kannry et al. 2020). However, the similarities between the Rogue, Eel, and Klamath rivers and close phylogenetic relationship among these populations suggest that this heterozygosity may be unique to the geographic region. Therefore, we caution that while a relatively large number of heterozygotes and absence of strong dominance may support a hopeful outlook for the maintenance early-migration associated alleles, the same pattern may not be true in other geographic regions.

We observe evidence of multigenerational interbreeding between fish bearing early- and late-migration haplotypes in the form of recombination among both late-summer run adults and half-pounders. Heterozygosity and recombination at migration timing markers might be reflective of historical conditions and maintained by a form of balancing selection such as temporally or spatially varying selection with gene flow (Thompson et al. 2020). Alternatively, our data might be providing a snapshot of a non-equilibrium process, wherein a recent, anthropogenic increase in connectivity between early and late migrators results in a temporary increase in the frequency of heterozygotes before swamping eliminates the less fit alleles (Thompson et al. 2019). The density of genotyping along the genome in our study is not great enough to thoroughly examine how many generations of interbreeding is responsible for the extent of recombination we observe. Ultimately, further research will be required to understand if the current level of gene flow between early and late migrators is reflective of historical conditions or is driven by relatively recent, anthropogenic changes in connectivity between early and late migrators.

**References**

Beechie T, Buhle E, Ruckelshaus M, Fullerton A, Holsinger L (2006) Hydrologic regime and the conservation of salmon life history diversity. Biological Conservation, 130, 560-572. https://doi.org/10.1016/j.biocon.2006.01.019

Boyle EA, Li YI, Pritchard JK (2017) An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell, 169, 1177-1186. https://doi.org/10.1016/j.cell.2017.05.038

Busby PJ, Wainright TC, Waples RS (1994) Status Review for Klamath Mountains Province Steelhead. National Oceanic and Atmospheric Administration, Technical Memorandum NMFS-NWFSC-19

Busby PJ, Wainright TC, Bryan GJ, Lierheimer LJ, Waples RS, Waknitz FW, Lagomarsino IV (1996) Status Review of West Coast Steelhead. National Oceanic and Atmospheric Administration, Technical Memorandum NMFS-NWFSC-27

Campbell NR, Harmon SA, Narum SR (2015) Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. Mol Ecol Resour, 15, 855-867. https://doi.org/10.1111/1755-0998.12357

Collins EE, Hargrove JS, Delomas TA, Narum SR (2020) Distribution of genetic variation underlying adult migration timing in steelhead of the Columbia River basin. Ecol Evol, 10, 9486-9502. https://doi.org/10.1002/ece3.6641

Coté C (2016) “Indigenizing” Food Sovereignty. Revitalizing Indigenous Food Practices and Ecological Knowledges in Canada and the United States. Humanities, 5. https://doi.org/10.3390/h5030057

Cramer (1985) Lost Creek Dam Fisheries Evaluation Phase I Completion Report: Summary and Recommendations Impacts of Lost Creek Dam on the Biology of Anadromous Salmonids in the Rogue River. US Army Corps of Engineers.

Crozier LG, McClure MM, Beechie T, Bograd SJ, Boughton DA, Carr M, Cooney TD, Dunham JB, Greene CM, Haltuch MA, Hazen EL, Holzer DM, Huff DD, Johnson RC, Jordan CE, Kaplan IC, Lindley ST, Mantua NJ, Moyle PB, Myers JM, Nelson MW, Spence BC, Weitkamp LA, Williams TH, Willis-Norton E (2019) Climate vulnerability assessment for Pacific salmon and steelhead in the California Current Large Marine Ecosystem. PLOS ONE, 14, e0217711. https://doi.org/10.1371/journal.pone.0217711

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol, 14, 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Evenson MD, Ewing RD (1992) Migration Characteristics and Hatchery Returns of Winter Steelhead Volitionally Released from Cole Rivers Hatchery, Oregon. North American Journal of Fisheries Management, 12, 736-743. https://doi.org/10.1577/1548-8675(1992)012<0736:Mcahro>2.3.Co;2

Everest FH (1971) An ecological and fish cultural study of summer steelhead in the Rogue River, Oregon. Federal Aid Progress Reports: Fisheries.

Everest FH (1973) Ecology and management of summer steelhead in the Rogue River. Federal Aid to Fish Restoration Project AFS31 Final Report.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 164, 1567-1587. https://doi.org/10.1093/genetics/164.4.1567

Ford, MN, Nichols K, Waples R, Anderson EC, Kardos M, Koch I, McKinney G, Miller MR, Myers J, Naish K, Narum S, O'Malley KG, Pearse D, Seamons T, Spidle A, Swanson P, Thompson TQ, Warheit K, Willis S. 2020. Reviewing and Synthesizing the State of the Science Regarding Associations between Adult Run Timing and Specific Genotypes in Chinook Salmon and Steelhead: Report of a workshop held in Seattle, Washington, 27–28 February 2020. U.S. Department of Commerce, NOAA Processed Report NMFS-NWFSC-PR-2020-06. https://doi.org/10.25923/mv3n-zb79

Greene CM, Hall JE, Guilbault KR, Quinn TP (2010) Improved viability of populations with diverse life-history portfolios. Biology Letters, 6, 382-386. https://doi.org/10.1098/rsbl.2009.0780

Groot C, Margolis L (1991) Pacific salmon life histories. UBC Press. Vancouver, British Columbia (Canada).

Hearsey JW, Kinziger AP (2015) Diversity in sympatric chinook salmon runs: timing, relative fat content and maturation. Environmental Biology of Fishes, 98, 413-423. https://doi.org/10.1007/s10641-014-0272-5

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. https://doi.org/10.1098/rspb.2015.3064

Hodge BW, Wilzbach MA, Duffy WG (2014) Potential Fitness Benefits of the Half-Pounder Life History in Klamath River Steelhead. Transactions of the American Fisheries Society, 143, 864-875. https://doi.org/10.1080/00028487.2014.892536

Hopelain JS (1998) Age, growth and life history of Klamath River Basin steelhead trout (Onchorhynchus mykiss irideus) as determined from scale analysis. California Department of Fish and Game Inland Fisheries Administrative Report, 98:3

Jacob C, McDaniels T, Hinch S (2010) Indigenous culture and adaptation to climate change: sockeye salmon and the St’át’imc people. Mitigation and Adaptation Strategies for Global Change, 15, 859-876. https://doi.org/10.1007/s11027-010-9244-z

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics, 27, 3070-3071. https://doi.org/10.1093/bioinformatics/btr521

Kannry SH, O'Rourke SM, Kelson SJ, Miller MR (2020) On the Ecology and Distribution of Steelhead (Oncorhynchus mykiss) in California's Eel River. J Hered, 111, 548-563. https://doi.org/10.1093/jhered/esaa043

Kardos M, Shafer ABA (2018) The Peril of Gene-Targeted Conservation. Trends Ecol Evol, 33, 827-839. https://doi.org/10.1016/j.tree.2018.08.011

Keefer ML, Wertheimer RH, Evans AF, Boggs CT, Peery CA (2008) Iteroparity in Columbia River summer-run steelhead (Oncorhynchus mykiss): implications for conservation. Canadian Journal of Fisheries and Aquatic Sciences, 65, 2592-2605. https://doi.org/10.1139/f08-160

Kesner WD, Barnhart RA (1972) Characteristics of the fall-run steelhead trout (Salmo gairdneri gairdneri) of the Klamath River system with emphasis on the half-pounder. California Fish and Game, 58, 204-220.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Molecular ecology resources. https://doi.org/10.1111/1755-0998.12387

Leider SA (1985) Precise Timing of Upstream Migrations by Repeat Steelhead Spawners. Transactions of the American Fisheries Society, 114, 906-908. https://doi.org/10.1577/1548-8659(1985)114<906:Ptoumb>2.0.Co;2

Leigh JW, Bryant D (2015) POPART: full-feature software for haplotype network construction. Methods in Ecology and Evolution, 6, 1110-1116. https://doi.org/10.1111/2041-210X.12410

McClure MM, Carlson SM, Beechie TJ, Pess GR, Jorgensen JC, Sogard SM, Sultan SE, Holzer DM, Travis J, Sanderson BL, Power ME, Carmichael RW (2008) Evolutionary consequences of habitat loss for Pacific anadromous salmonids. Evol Appl, 1, 300-318. https://doi.org/10.1111/j.1752-4571.2008.00030.x

McPhee MV, Utter F, Stanford JA, Kuzishchin KV, Savvaitova KA, Pavlov DS, Allendorf FW (2007) Population structure and partial anadromy in Oncorhynchus mykiss from Kamchatka: relevance for conservation strategies around the Pacific Rim. Ecology of Freshwater Fish, 16, 539-547. https://doi.org/10.1111/j.1600-0633.2007.00248.x

Merilä J, Hendry AP (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evolutionary Applications, 7, 1-14. https://doi.org/10.1111/eva.12137

Micheletti SJ, Hess JE, Zendt JS, Narum SR (2018a) Selection at a genomic region of major effect is responsible for evolution of complex life histories in anadromous steelhead. BMC Evol Biol, 18, 140. https://doi.org/10.1186/s12862-018-1255-5

Micheletti SJ, Matala AR, Matala AP, Narum SR (2018b) Landscape features along migratory routes influence adaptive genomic variation in anadromous steelhead (Oncorhynchus mykiss). Mol Ecol, 27, 128-145. https://doi.org/10.1111/mec.14407

Mimura M, Yahara T, Faith DP, Vázquez‐Domínguez E, Colautti RI, Araki H, Javadi F, Núñez‐Farfán J, Mori AS, Zhou S (2017) Understanding and monitoring the consequences of human impacts on intraspecific variation. Evolutionary applications, 10, 121-139. https://doi.org/10.1111/eva.12436

Nesbitt HK, Moore JW, Manning P (2016) Species and population diversity in Pacific salmon fisheries underpin indigenous food security. Journal of Applied Ecology, 53, 1489-1499. https://doi.org/10.1111/1365-2664.12717

ODFW (1990) Effects of Lost Creek Dam on winter steelhead in the Rogue River, phase II. ODFW, Fish Research Project DACW 57-77-C-0033, Completion Report.

ODFW (1994) Effects of Lost Creek Dam on summer steelhead in the Rogue River, phase II. ODFW, Fish Research Project DACW 57-77-C-0033, Completion Report.

Oke KB, Hendry AP (2019) Genetic insights into the past, present, and future of a keystone species. Proc Natl Acad Sci U S A, 116, 344-346. https://doi.org/10.1073/pnas.1819789116

Oomen RA, Kuparinen A, Hutchings JA (2020) Consequences of Single-Locus and Tightly Linked Genomic Architectures for Evolutionary Responses to Environmental Change. J Hered, 111, 319-332. https://doi.org/10.1093/jhered/esaa020

Paradis E (2010) pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics, 26, 419-420. https://doi.org/10.1093/bioinformatics/btp696

Pascual M, Bentzen P, Riva Rossi C, Mackey G, Kinnison MT, Walker R (2001) First Documented Case of Anadromy in a Population of Introduced Rainbow Trout in Patagonia, Argentina. Transactions of the American Fisheries Society, 130, 53-67. https://doi.org/10.1577/1548-8659(2001)130<0053:Fdcoai>2.0.Co;2

Pearse DE, Garza JC, Myers J, Spence B (2019) Northern California steelhead DPS-Configuration Review-Panel Report media.fisheries.noaa.gov/dam-migration/508compliant\_pearse\_etal\_2019\_ncsteelhead\_petitionpanelfinalreport\_9\_30\_2019. pdf. Accessed August 2021

Pearse DE, Miller MR, Abadia-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. https://doi.org/10.1098/rspb.2014.0012

Peterson ML, Hankin DG, Manishin K (2017) Decline in the Half-Pounder Life History among Trinity River, California, Steelhead. Transactions of the American Fisheries Society, 146, 1245-1261. https://doi.org/10.1080/00028487.2017.1356372

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. https://doi.org/10.1126/sciadv.1603198

Quinn TP, McGinnity P, Reed TE, Bradford M (2016) The paradox of “premature migration” by adult anadromous salmonid fishes: patterns and hypotheses. Canadian Journal of Fisheries and Aquatic Sciences, 73, 1015-1030. https://doi.org/10.1139/cjfas-2015-0345

Reisenbichler RR, McIntyre JD, Solazzi MF, Landino SW (1992) Genetic Variation in Steelhead of Oregon and Northern California. Transactions of the American Fisheries Society, 121, 158-169. https://doi.org/10.1577/1548-8659(1992)121<0158:Gvisoo>2.3.Co;2

Rivers CM (1963) Rogue River Fisheries, Volume 1: History and development of the

Rogue River Basin as related to its fishery prior to 1941. Oregon State Game Comm.

Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. Evolution, 66, 1-17. https://doi.org/10.1111/j.1558-5646.2011.01486.x

Roelofs (1983) Current status of California summer steelhead (Salmo gairdneri) stocks

and habitat, and recommendations for their management. Submitted to USDA Forest

Service, Region 5. https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=32669 Accessed May 2021

Satterthwaite WH, Carlson SM (2015) Weakening portfolio effect strength in a hatchery-supplemented Chinook salmon population complex. Canadian Journal of Fisheries and Aquatic Sciences, 72, 1860-1875. https://doi.org/10.1139/cjfas-2015-0169

Scheet P, Stephens M (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. The American Journal of Human Genetics, 78, 629-644. https://doi.org/10.1086/502802

Schindler DE, Armstrong JB, Reed TE (2015) The portfolio concept in ecology and evolution. Frontiers in Ecology and the Environment, 13, 257-263. https://doi.org/10.1890/140275

Schindler DE, Hilborn R, Chasco B, Boatright CP, Quinn TP, Rogers LA, Webster MS (2010) Population diversity and the portfolio effect in an exploited species. Nature, 465, 609-612. https://doi.org/10.1038/nature09060

Sheer MB, Steel EA (2011) Lost Watersheds: Barriers, Aquatic Habitat Connectivity, and Salmon Persistence in the Willamette and Lower Columbia River Basins. Transactions of the American Fisheries Society, 135, 1654-1669. https://doi.org/10.1577/t05-221.1

Swezey SLH, R. F. (1977) Ritual Management of Salmonid Fish Resources in California. The Journal of California Anthropology, 4.

Thioulouse J, Dray S (2007) Interactive multivariate data analysis in R with the ade4 and ade4TkGUI packages. Journal of Statistical Software, 22, 1-14. https://doi.org/10.18637/jss.v022.i05

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. https://doi.org/10.1126/science.aba9059

Thompson TQ, Bellinger MR, O'Rourke SM, Prince DJ, Stevenson AE, Rodrigues AT, Sloat MR, Speller CF, Yang DY, Butler VL, Banks MA, Miller MR (2019) Anthropogenic habitat alteration leads to rapid loss of adaptive variation and restoration potential in wild salmon populations. Proc Natl Acad Sci U S A, 116, 177-186. https://doi.org/10.1073/pnas.1811559115

Waples RS, Ford MJ, Nichols K, Kardos M, Myers J, Thompson TQ, Anderson EC, Koch IJ, McKinney G, Miller MR, Naish K, Narum SR, O’Malley KG, Pearse DE, Pess GR, Quinn TP, Seamons TR, Spidle A, Warheit KI, Willis SC, Baker CS (2022) Implications of Large-Effect Loci for Conservation: A Review and Case Study with Pacific Salmon. Journal of Heredity, 113, 121-144. https://doi.org/10.1093/jhered/esab069

Waples RS, Lindley ST (2018) Genomics and conservation units: The genetic basis of adult migration timing in Pacific salmonids. Evol Appl, 11, 1518-1526. https://doi.org/10.1111/eva.12687

Warren DR, McClure MM (2012) Quantifying Salmon-Derived Nutrient Loads from the Mortality of Hatchery-Origin Juvenile Chinook Salmon in the Snake River Basin. Transactions of the American Fisheries Society, 141, 1287-1294. https://doi.org/10.1080/00028487.2012.686950

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. evolution, 1358-1370.

Willis SC, Hess JE, Fryer JK, Whiteaker JM, Brun C, Gerstenberger R, Narum SR (2020) Steelhead (Oncorhynchus mykiss) lineages and sexes show variable patterns of association of adult migration timing and age‐at‐maturity traits with two genomic regions. Evolutionary Applications. https://doi.org/10.1111/eva.13088

**Statements and Declarations**

The authors declare that no funds, grants, or other support were received to assist with the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose. Detailed logs and all data needed to replicate this study are available at a github repository: <https://github.com/david-dayan/half-pounder> archived at zenodo with a stable identifier ([DOI:10.5281/zenodo.4750859](https://zenodo.org/badge/latestdoi/302479383)). Raw sequence data is archived at the short read archive with BioProject accession number PRJNA793293.

**Tables**

**Table 1:** Rogue River steelhead migration timing nomenclature. The peak of freshwater entry timing and spawning for each run occurs sometime during the periods defined, but is much narrower in a given year.

|  |  |  |  |
| --- | --- | --- | --- |
| **Run** | **Peak Freshwater Entry Time\*** | | **Spawn time\*** |
| Late-Summer | August-October | December-March | |
| Half-Pounder | August-October | December-March\*\* | |
| Early-Summer | May-July | December-March | |
| Winter | November-March | March-June | |

\*Ranges as reported in (Everest 1971, 1973; ODFW 1990)

\*\*While half-pounders have not been reported as present in a winter steelhead survey (ODFW 1990), half-pounders may be present on spawning grounds with winter because they maintain fertility longer than adults (Everest 1973), and previous surveys during this time period have not explicitly searched for half-pounders.

**Table 2:** Sample sizes in the final filtered dataset, nHOR and nNOR refer to sample sizes of hatchery-origin and natural origin, respectively. nXX , nXY and n00, refer to number of female, male and undetermined sex genotypes observed. Year refers to sampling time, not brood year.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Run** | **Year** | **n** | **nHOR:NOR** | **nXX:XY:00** | **Sampling Dates** | **Location** |
| Late-Summer | 2019 | 157 | 0:157 | 94:60:3 | July 26th –  October 21st | Huntley Park (rkm 13) |
| Late-Summer | 2020 | 118 | 0:118 | 71:45:2 | July 29th –  October 16th | Huntley Park (rkm 13) |
| Half-Pounder | 2018 | 338 | 0:338 | 173:160:5 | September 7th –  October 1st | Huntley Park (rkm 13) |
| Half-Pounder | 2019 | 305 | 0:305 | 171:131:3 | August 14th –  September 25th | Huntley Park (rkm 13) |
| Early-Summer | 2019 | 42 | 41:1 | 25:17:0 | June 26th | Cole M. Rivers Hatchery (rkm 253) |
| Winter | 2019 | 40 | 18:22 | 24:16:0 | May 1st -  May 13th | Cole M. Rivers Hatchery (rkm 253) and Applegate Dam (rkm 154 + 75) |

**Table 3:** Pairwise differentiation (Weir and Cockerham FST) among the three adult runs and half-pounder run. FST estimates using the neutral dataset markers are below the diagonal, estimates using the migration timing dataset are above the diagonal.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Late-Summer** | **Half-Pounder** | **Early-Summer** | **Winter** |
| **Late-Summer** |  | 0.0000 | 0.4419 | 0.2401 |
| **Half-Pounder** | 0.0002 |  | 0.4325 | 0.2094 |
| **Early-Summer** | 0.0105 | 0.0087 |  | 0.9254 |
| **Winter** | 0.0045 | 0.0034 | 0.0104 |  |

**Table 4:** Number of late-summer and half-pounder individuals assigned as genetically similar to early migrators, late migrators or unassigned. Individuals are assigned as genetically similar to early or late migrators if their individual LD1 scores fall within the LD1 intervals observed among with early-summer or winter runs.

|  |  |  |  |
| --- | --- | --- | --- |
| **Run** | **Late-Migrator Assigned** | **Early-Migrator Assigned** | **Unassigned** |
| **Late-Summer** | 57 (21%) | 3 (1%) | 215 (78%) |
| **Half-Pounder** | 228 (35%) | 76 (12%) | 339 (53%) |

**Figures**

**Map

Description automatically generated**

**Fig. 1** Sampling locations along the Rogue River, Oregon, USA.

From the editor: “Maps for international journals need to be geo-referenced with either lat/long coordinations or a continent-level inset map. Subnational political jurisdictions should not be assumed. “

We need a new or revised map figure. I’m still working on this, but will likely just add latitude and longitude manually to the existing figure using some landmarks.

**A picture containing text, circle, diagram, screenshot

Description automatically generated**

**Fig. 2:** PCA using the neutral dataset, individuals are colored according to run after ordination.

A picture containing screenshot, colorfulness, graphic design, graphics

Description automatically generated

**Fig. 3** Admixture proportions from STRUCTURE for *k =* 2 (top) to 4 (bottom) ~~for the four runs~~ using the neutral dataset. The full STRUCTURE results have been randomly down sampled to 42 individuals per run. Results for all individuals available in Supplemental Fig. 1. Collections drawn from multiple years are combined.

**A picture containing screenshot, text, diagram, plot

Description automatically generated**

**Fig. 4:** Distribution of individual discriminant analysis 1 scores (LD1 Score) from *a priori* DAPC trained on only early-summer and winter collections and the migration timing dataset. LD1 scores were used to classify late-summer and half-pounder steelhead as genetically early or late migrators. Bars are colored by run. Collections drawn from multiple years are combined.

**A picture containing text, diagram, screenshot, line

Description automatically generated**

**Fig. 5:** Late-summer run individual LD1 Score vs. Julian day of sampling at river kilometer 13. Line is the fitted value of Julian day of sampling and its 95% confidence interval on the y-axis *vs.* LD1 score from a linear model. Assignment bins from the DAPC (range of LD1 scores for known early-summer and winter individuals) are shown in yellow (winter) and green (early-summer). Hollow and filled circles represent individuals from 2019 and 2020, respectively.

Chart

Description automatically generated

**Fig. 6** Allele frequency among the three adult and half-pounder runs for migration timing dataset. Dendrograms reflect hierarchical clustering of allele frequencies across all groups and markers. Collections drawn from multiple years are combined.

A yellow and purple squares

Description automatically generated with low confidence

**Fig. 7** Haplotypes near the *greb1l-rock1* region on chromosome 28 inferred from phased migration timing informative markers. Each horizontal line represents a single haplotype (each individual is represented twice in the figure). Haplotypes from late-summer run and half-pounder individuals assigned as early and late migrators are identical to early-summer and winter run results and are not displayed. Haplotypes are hierarchically clustered within each run. Alleles are polarized such that purple (dark) represents early-migration associated allele and yellow (light) represents a late-migrator associated allele. Markers are ordered according to their position on chromosome 28. Collections drawn from multiple years are combined.

**Supplemental Tables**

**Supplemental Table 1**: Observed *vs.* Expected heterozygosity for neutral and migration timing datasets among runs. Collections drawn from multiple years are combined.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Winter** | | **Early-Summer** | | **Late-Summer** | | **Half-Pounder** | |
|  | Ho | He | Ho | He | Ho | He | Ho | He |
| **Neutral Dataset** | ﻿0.305 | ﻿0.311 | 0.301 | 0.304 | 0.301 | 0.308 | 0.299 | 0.309 |
| **Migration Timing**  **Dataset** | 0.102 | 0.096 | 0.010 | 0.009 | 0.387 | 0.355 | 0.266 | 0.357 |

**Supplemental Table 2:** Meanadmixture proportions and standard deviation for the four runs analyzed with STRUCTURE using *k = ­*3. Cluster 1 appears as purple, 2 as blue-green, and 3 as yellow (in order from top to bottom) in the *k* = 3 row of Fig. 2

|  |  |  |  |
| --- | --- | --- | --- |
| **Run** | **Cluster 1** | **Cluster 2** | **Cluster 3** |
| Late-Summer | 0.37 ± 0.18 | 0.33 ± 0.21 | 0.30 ± 0.16 |
| Half-Pounder | 0.34 ± 0.17 | 0.32 ± 0.21 | 0.34 ± 0.18 |
| Early-Summer | 0.32 ± 0.17 | 0.17 ± 0.09 | 0.51 ± 0.18 |
| Winter | 0.41 ± 0.18 | 0.18 ± 0.08 | 0.42 ± 0.18 |

**Supplemental Table 3:** Meanadmixture proportions and standard deviation for the four runs analyzed with STRUCTURE using *k = ­*4. Cluster 1 appears as purple, 2 as blue, 3 as green, and 4 as yellow (in order top to bottom) in the *k* = 3 row of Fig. 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Run** | **Cluster 1** | **Cluster 2** | **Cluster 3** | **Cluster 4** |
| Late-Summer | 0.27 ± 0.16 | 0.25 ± 0.23 | 0.22 ± 0.17 | 0.26 ± 0.13 |
| Half-Pounder | 0.25 ± 0.15 | 0.24 ± 0.24 | 0.25 ± 0.20 | 0.26 ± 0.14 |
| Early-Summer | 0.27 ± 0.17 | 0.09 ± 0.06 | 0.42 ± 0.24 | 0.21 ± 0.12 |
| Winter | 0.29 ± 0.16 | 0.09 ± 0.05 | 0.31 ± 0.20 | 0.31 ± 0.15 |

**Supplemental Figure Captions**

**A picture containing screenshot, colorfulness, art

Description automatically generated**

**Supplemental Fig. 1:** Admixture proportions from STRUCTURE for *k =* 2 – 4 for each of the four runs. Collections drawn from multiple years are combined.

A picture containing text, screenshot, diagram, design

Description automatically generated

**Supplemental Fig. 2:** Linkage disequilibrium (r2) among migration timing markers in (a) winter and early-summer run, (b) late-summer run and half pounders.

Chart

Description automatically generated with low confidence

**Supplemental Fig. 3** Haplotype network. Each node represents a unique haplotype observed in the dataset. Colors within nodes represent proportion of each run with that haplotype. Edges represent the inferred minimum spanning tree connecting haplotypes. Collections drawn from multiple years are combined.

**Online Resources**

**Online Resource 1** Genotyping Notebook. This html file can be opened with a web browser and contains narrative logs of all work completed to produce filtered genotypes from raw sequencing reads with integrated code, results and commentary.

**Online Resource 2** Analysis Notebook. This html file can be opened with a web browser and contains narrative logs of all analyses along with integrated code, results and commentary.

**Online Resource 3** Metadata. This excel spreadsheet contains all metadata used in the analysis.

**Online Resource 4** GT-seq Panel Info. This excel spreadsheet contains read mapping results for the SFGL *O. mykiss* GT-seq Panel, as well as whether each marker passed genotype quality filter and is included in either the neutral or the migration timing datasets.