Title: A major effect locus involved in migration timing is shared by pink and sockeye salmon

Key words: *lrrc9*, complex phenotype, local adaptation, life history diversity, low coverage whole genome sequencing

Authors: Patrick D. Barry[1], David A. Tallmon[2], Natasha S. Howe[3], Diana S. Baetscher[1], Katie L. D'Amelio[1], Scott C. Vulstek[1], Josh R. Russell[1], Anthony J. Gharrett[4], & Wesley A. Larson[1]

[1] NOAA Alaska Fisheries Science Center, Auke Bay Laboratories, Juneau, Alaska, USA
[2] University of Alaska Southeast, Biology and Marine Biology Program, Juneau, Alaska, USA
[3] Pacific States Marine Fisheries Commission, Seattle, Washington, USA
[4] University of Alaska Fairbanks, College of Fisheries and Ocean Science, Juneau, Alaska, USA

ABSTRACT

Migration is an important component of the life cycles of many organisms and differences in migration timing can greatly influence fitness. Variation in migration timing (hereafter referred to as run timing) is found in many salmon species and contributes to the portfolio of life history diversity that is vital for maintaining healthy populations. Pink salmon (*Oncorhynchus gorbuscha*) are composed of temporally isolated lineages that spawn in even and odd years. Here, we focus on early and late runs of pink salmon in Auke Creek, Alaska, which spawn a few weeks apart in the same ≈50m stretch of stream. We conducted low coverage, whole genome sequencing on pink salmon from the early and late runs of both even and odd lineages to investigate the genetic basis of run timing in this system. Within the even lineage we detected a major effect locus displaying high genetic divergence between early and late run individuals on chromosome 10 of the pink salmon genome. This same locus showed a lower - but still elevated - amount of genetic divergence in the odd lineage. The region of high divergence includes the Leucine Rich Repeat Containing 9 gene (*lrrc9*), and genetic divergence was substantially reduced outside of this gene. The *lrrc9* gene also shows high divergence between two ecotypes of sockeye salmon (*Oncorhynchus nerka*) that differ in both run timing and spawning habitat. Comparison of the *lrrc9* gene region in pink and sockeye salmon shows that variation within this gene is associated with run timing in both species. However, the polymorphic sites (i.e, SNPs) and run timing alleles are unique to each species and arose from independent mutations following divergence from a common ancestor. These results demonstrate that adaptation in pink salmon can occur at small temporal scales and that complex traits can be controlled by similar genetic mechanisms in multiple species.

Introduction

Large-scale animal migrations are iconic seasonal events undertaken by a diverse array of taxa, including insects, birds, reptiles, large mammals, and fishes (Dingle 2017). Annual migration of massive numbers of salmon from marine to freshwater ecosystems to spawn influences the behavior, diet, and distribution of a variety of taxa that consume salmon (Willson & Halupka, Hildebrand et al. 1999, Gende et al. 2003, Schindler et al 2013, Levi et al 2015), including humans (Campbell and Butler 2010, Halffman et al. 2015).

Biotic and abiotic conditions vary dramatically across the range of watersheds used for spawning by wild Pacific salmon throughout the Pacific Rim. Strong natural selection on traits closely associated with reproduction has led to fine-scale adaptation to local environmental conditions and a diverse array of phenotypes and life histories among populations (Ricker 1972, Garcia de Leaniz 2007, Fraser et al. 2011). One of the most prominent traits associated with salmon reproduction is the timing of annual adult spawning migration ('run timing' hereafter) to freshwater, which is thought to be optimized based on environmental conditions, such as stream temperature and flow, to maximize reproductive success (Beacham and Murrary 1990, Smoker et al. 1998, Eliason et al. 2011, Gharrett et al. 2013, Manhard et al. 2018, Koch and Narum 2021).

As a result of the link between salmon run timing and reproductive success, there is great interest in understanding the genetic basis of run timing. Quantitative genetic studies show considerable additive genetic variance and heritability of run timing (h^2~0.5; Carlson and Seamons 2008), although the amounts differ among species and populations (Smoker et al. 1998, Carlson and Seamons 2008, Manhard 2018, Koch and Narum 2021). Until recently it was hypothesized that run timing was controlled by many genes of small effect (Waples et al. 2022). Recent genomic studies have identified some individual loci or large chromosomal regions that strongly influence run timing of Atlantic salmon (*Salmo salar*; Cauweiler et al 2018) and Chinook salmon (*Oncorhynchus tshawytscha*) and steelhead trout (*Oncorhynchus mykiss;* O'Malley et al. 2010, Hess et al 2016, Prince et al. 2017, Pearse 2019, Thompson et al. 2019). For example, genotypes within the *greb1l/rock1* region explain a substantial amount of variation in run timing of both Chinook salmon (67-85%; Thompson et al. 2019) and steelhead trout (36-43%; Narum et al. 2023). Whole genome sequencing data have facilitated detailed investigation of the *greb1l/rock1* region and revealed that (1) alleles are more similar within than among species, suggesting independent selection on these genes, (2) the region does not appear to be a structural variant (e.g., chromosomal inversion), and (3) the region contains multiple potential duplications, which may influence its function (Micheletti et al. 2018, Horn and Narum 2023, Waples et al. 2022).

Genomic studies of salmon run timing have often been conducted on species and populations with discrete runs that vary seasonally - differences measured in months. For example, the median date of Chinook salmon spring and fall runs in California is separated by at least three months (Thompson et al. 2020), and show neutral genetic divergence (F_{ST} > 0.03, Prince et al. 2017). However, ecologically important run timing differences that span smaller timescales (e.g., weeks) also exist in other salmon species. Sockeye salmon (*Oncorhynchus nerka*) populations that spawn on lake beaches often spawn just a few weeks later than populations that spawn in nearby inlet streams, likely due to thermal differences between habitats (Hilborn et al. 2003).

This run timing variation represents a bet hedging strategy that ensures catastrophic events such as anomalously warm water temperatures do not affect the entire run and also increases the timeframe over which spawning salmon serve as an ecosystem service to be consumed by predators (Schindler et al. 2013).

Pink salmon (*Oncorhynchus gorbuscha*), which are the most abundant Pacific salmon species, also display variation in run timing across small spatial scales, likely reflecting fine-scale environmental variation (Smoker et al. 1998, Gharrett et al. 2013). Pink salmon are particularly interesting because they frequently occur in the same streams as even- and odd-year lineages, which are reproductively isolated from one another by their strict two year life cycle (Heard 1991). In other words, pink salmon lineages that spawn in a given stream in even and odd years do not naturally have gene flow between them and are evolutionarily independent, yet are subjected to the same selective pressures imposed by local abiotic conditions over long periods of time (Oke et al. 2019). Reproductive isolation between the odd- and even-year lineages is so complete that pink salmon from the two lineages in the same stream are more genetically divergent than pink salmon from the same lineage that spawn on different continents (Hawkins et al. 2002, Tarpey et al. 2018). Odd year populations have higher levels of genetic variation than even year populations, and a number of chromosomal regions differ between the two lineages, including regions that could underlie parallel adaptations (Limborg et al. 2014, Seeb et al. 2014, Tarpey et al. 2018, Christensen et al. 2021).

Long-term molecular surveys and experimental studies of pink salmon in Auke Creek, Alaska, highlight adaptive differences between odd- and even-year lineages, despite shared selection pressures. Although both odd and even year fish spawn primarily in the same ~50 m stretch of Auke Creek, second generation hybrids created by experimentally crossing odd and even year fish exhibited outbreeding depression in marine survival not seen in first generation hybrids or controls, demonstrating the evolution of different coadapted gene complexes in each lineage (Gharrett and Smoker 1991, Gharrett and Smoker 1999). However, both median juvenile outmigration and adult run timing have advanced in both lineages since the 1970s, as stream temperatures have increased and late arriving individuals have been selected against (Fig. 1; Taylor et al. 2008, Kovach et al. 2012, Kovach 2013a, Manhard et al. 2017).

Fine-scale adaptive variation within the odd- and even-year lineages of Auke Creek pink salmon has also been documented. Both odd- and even-year spawners arrive at Auke Creek in early and late runs, the modes of which are separated by ~20 days (Taylor et al. 1980, Smoker et al. 1998, Manhard et al. 2017). Experimental crosses revealed additive genetic variation underlying both adaptive differences in run timing (Smoker et al. 1998, Manhard 2018) and juvenile developmental rates of early and late run fish (Hebert et al. 1998, Echave et al. 2017, Gharrett et al. 2013). These adaptive differences evolved despite the homogenizing effects of considerable gene flow between early and late runs within lineages inferred from low levels of divergence at allozyme and microsatellite marker loci (MacGregor et al. 1998, Kovach et al. 2013b).

We investigated the genomic architecture of run timing in odd- and even-year Auke Creek pink salmon using low-coverage whole genome sequencing. Given the importance of run timing to reproductive success and recent shifts in Auke Creek pink salmon migration timing, we aimed to

identify loci associated with adaptive differences between early- and late-run pink salmon in both lineages. We genotyped early and late migrating salmon from both odd- and even-year lineages and identified a region of the genome showing large genetic divergence associated with run timing. We then compared the variation observed within pink salmon with two ecotypes of sockeye salmon that both spawn in different habitats and differ in migration timing, demonstrating that this region includes loci implicated in run timing differences between beach and tributary sockeye salmon ecotypes (Veale & Russello 2017). Our results illustrate that large-effect loci can facilitate fine-scale adaptation in salmon (run timing separated by weeks, spawning locations 10s of m apart), even when neutral genetic divergence throughout the genome is low. These findings highlight the importance of understanding and conserving fine-scale diversity to ensure that salmon populations are resilient to future challenges.

Methods

Study system and sample collection

The Auke watershed is composed of multiple small tributaries that feed into Auke Lake, a small lake in Southeast Alaska. Auke Creek is a short (≈350 m) and moderately steep (2–4% gradient) outlet stream connecting Auke Lake to the ocean (Fig 1). Stream discharge, which peaks in September but varies annually with precipitation and snowpack, influences the frequency and intensity of pink salmon ascending Auke Creek. The majority of pink salmon spawn within a small section (≈50 m) of Auke Creek below the lake and above the head of tidewater (Fig. 1), while limited spawning occurs within the intertidal and in Lake Creek (Manhard et al. 2017, Taylor 2008, Fukushima & Smoker 1997).

Since 1980, the National Marine Fisheries Service has operated a research station, consisting of a permanent weir and experimental hatchery, located above the high tide line. Daily counts of pink salmon have been taken during juvenile spring and adult fall migrations. Each fish is hand-passed across the weir providing precise return date paired with daily environmental data.

For detailed laboratory and analytical methods please refer to the supplementary methods document. Below we briefly describe these methods.

Sample selection and whole genome sequencing

Pink salmon samples selected for whole genome sequencing were collected from 2001-2006 originally for experimental crosses to evaluate outbreeding depression between early and late spawners (Manhard et al. 2018). We selected 288 Auke creek pink salmon for sequencing, targeting equal representation between run-timing, lineage, and sex (Table 1). Sequencing libraries were prepared following Baym et al. (2015) and Therkildsen and Palumbi (2017). Samples were sent to Novogene (Sacramento, CA) for whole genome sequencing with paired-end 150-bp reads on an Illumina NovaSeq.

Analysis of whole genome data in pink salmon

Sequencing files were trimmed of adapters with TRIMMOMATIC v0.39 (Bolger, Lohse, and Usadel 2014) and poly-G tails with the program fastp (Chen et al. 2018). Trimmed sequences were aligned to the pink salmon reference genome (GCA_021184085.1_OgorEven_v1.0_genomic.fna) with BWA. SAMtools (Li et al. 2009) was used to convert between file formats. PCR duplicates were removed with picard

(http://broadinstitute.github.io/picard/), while BamUtil (Jun et al. 2015) was used to clip overlapping read pairs. Sequencing depth for each individual was calculated with SAMtools (Li et al. 2009) and samples whose sequencing coverage fell below a mean of 1x were dropped from analyses.

Genotype likelihoods (GLs) and pairwise F_{ST} were calculated with ANGSD v0.933 (Korneliussen et al., 2014) using the SAMtools genotype likelihood calling algorithm. Sample GLs were used to estimate individual allele frequencies and a covariance matrix with PCAngsd (Meisner & Albrechtsen 2018) for the whole genome and specific regions of the genome displaying high divergence between runs. Eigendecomposition of the covariance matrix was performed with R statistical software (v.4.3.2; R Core Team 2021) and visualized with ggplot2 (Wickham 2016). Pairwise linkage disequilibrium (LD measured as r^2) was estimated with GLs in ngsLD v1.2.0 (Fox et al. 2019) for a 300 Kb region of chromosome 10, which displayed highly elevated differentiation. We also generated a genotype heatmap across all individuals for loci within this region. Fractional allele dosages were calculated for all individuals after polarizing genotype likelihoods to a reference individual. Polarization effectively sets the major allele to the chosen reference individual to compare similar genotypes across loci. We used an early-returning even-lineage female pink salmon as reference. Loci for which individuals had flat genotype likelihoods, indicating no sequencing read depth, were omitted. Fractional allele dosages were then rounded and visualized with the ComplexHeatmap package in R (Gu 2022).

To evaluate the effect of SNPs within the *lrrc9* gene on migration timing, individuals were assigned to the early, heterozygous, and late alleles based on PCA score and SNP heatmap pattern, and pairwise F_{ST} was calculated among each group as above. Coding sequence (cds) locations for the *lrrc9* gene were identified and extracted from the annotated reference genome. The SNPs differentiating the early and late allele ($F_{ST} \geq 0.8$) were then substituted into the codings sequence and translated in Geneious Prime 2023.2.1 to subset on possible functional mutations. The resulting proteins were then compared against the protein product from the pink salmon reference sequence (XP_046220528.1).

Comparing patterns of variation across species

To investigate if a large-effect locus associated with run timing in pink salmon (see results) was also present in sockeye salmon, data from Euclide et al. (2023) were re-analyzed and compared to Auke Creek pink salmon. The data from Euclide et al. (2023) consisted of sockeye salmon samples from seven populations in Bristol Bay, Alaska, USA sequenced to an average depth of 6.6x (see Euclide et al., 2023 for sequencing methods). These individuals were primarily sampled from beach and tributary populations, which diverge at a region on chromosome 12 containing the *lrrc9* gene (Nichols et al., 2016; Veale & Russello, 2017; Euclide et al., 2023).

A subset of the pink salmon samples from this study, and a subset of the Euclide et al. (2023) sockeye samples (Table S1), were aligned to a phylogenetically intermediate species, chum salmon (*Oncorhynchus keta*; Alexandrou et al., 2013; Kitano et al., 1997). The pink salmon subset consisted of ten samples from each lineage and run timing (even-early, even-late, odd-early, and odd-late) with the greatest depth of coverage for a total of 40 pink salmon samples. Sockeye salmon data consisted of four individuals from each of seven sampling locations, which included both beach and tributary ecotypes (28 total, Euclide et al., 2023). Sequencing files

were processed as above for pink salmon; however, to reduce potential batch effects between the pink and sockeye salmon sequencing runs the sockeye sequencing files were down-sampled and depth thresholds were adjusted (see supplementary methods for details).

Analyses comparing patterns across species focused on variation at the *lrrc9* gene on chromosome 35 of the chum salmon genome from 28.1 to 28.2 Mb. PCA clusters within each species were used to assign putative genotypes to individuals. For each species, pairwise F_{ST} was calculated between homozygous genotypes. Within species F_{ST} values within the *lrrc9* region were compared to determine if the same loci were differentiated in both species.

Exon and cds locations were determined from the annotated chromosome 35 from the chum salmon genome and compared to SNP positions in the region of interest. Again, to determine if mutations within the *lrrc9* region were synonymous or non-synonymous, the SNPs identified within sockeye and pink salmon were substituted into the coding sequence and translated in Geneious Prime 2023.2.1. The resulting protein was compared against the protein product from the reference sequence (XP_052352787).

Results:

Sequencing statistics and quality control

A total of 288 individuals were sequenced. After removing 12 individuals that had low average read depth (<1x), we retained 276 individuals with an average depth of 2.05x (0.56 SD). There were 6.7 million SNPs, 5.2 million in the even lineage, 4.8 million in the odd lineage and 3.26 million shared between lineages.

Population structure and genome-wide patterns of divergence

As expected, even- and odd-year lineages of pink salmon in Auke Creek, Alaska were genetically divergent from one another ($F_{ST\ weighted}$ = 0.07; Fig. S1). Most individuals grouped tightly by lineage in the PCA with 63% of the variance explained by PC1 separating the even and odd lineages. Two outliers, both late returning females, were identified in the odd lineage (Fig. S1) which were separated from the main group on PC2. These two individuals clustered away from the main group of odd lineage individuals when a PCA was conducted on only the odd year lineage samples; however only 2.16% of the variation was explained along this axis. While there was substantial overlap of early and late individuals in the odd lineage PCA, there was separation of early and late individuals in the even lineage along the first principal component with 1.57% of the variation explained (Fig S1).

Even- and odd-year lineages demonstrated markedly different patterns of genome-wide divergence between early and late runs. The highest divergence between early and late runs was in the even-year lineage on Chromosome 10 between 72.70 and 72.74 Mb ($\bar{F}_{ST}$ = 0.52, SD = 0.16), in marked contrast to the low divergence across the rest of the genome ($\bar{F}_{ST}$ = 0, SD = 0.01; Fig. 2A). Variation in this region did not differ by sex, as sex-specific genome scans of the even lineage exhibited a consistent peak at the same location on chromosome 10, albeit with slightly different magnitudes, among all comparisons between runs (Fig. S2). Within the even year lineage there are two additional peaks of smaller magnitude on chromosome 17 ($\bar{F}_{ST}$ = 0.13 and 0.22 respectively), the larger of which spans 23.2kb. Within the odd lineage, genome-wide divergence between early and late runs was small ($\bar{F}_{ST}$ = 0, SD = 0.01) and lacked obvious

corresponding peaks on Chromosomes 10 and 17 present in the even lineage (Fig. 2B). There were, however two small peaks identified in a 8.4kb region in the middle of Chromosome 3 ($\bar{F}_{ST}$ = 0.04, SD = 0.05) and a 2.3kb region on Chromosome 8 ($\bar{F}_{ST}$ = 0.14, SD = 0.09).

Characterizing Chromosome 10 region

We further investigated the highly diverged 300kb region (72.60–72.90 Mb) of chromosome 10, the region of strongest association (RoSA) with run timing. Within the RoSA there were 779 SNPs identified in the even- and odd-year lineages, 464 of which were shared. There were 8 genes identified within this region, seven protein coding and one pseudo gene (Fig. 3, Table 2). The region of highest differentiation in the even lineage tightly overlaps the leucine-rich repeat containing 9 gene (*lrrc9*; Fig. 3A). Lower but still elevated divergence in the even lineage was also observed within the reticulon 1a gene (*rtn1a*) and upstream of the six homeobox 6a gene (*six6a*; Fig. 3A). The *lrrc9* gene was also slightly diverged between runs in the odd lineage ($\bar{F}_{ST}$ = 0.03, SD = 0.03).

Principal component analysis restricted to the plateau of high F_{ST} values in the even lineage found between 72.70–72.74 Mb on chromosome 10 displayed three discrete groups of individuals (Fig. 3B). A SNP heatmap, calculated from the rounded allele dosage using an early individual from the even lineage as a reference, revealed consistent differences among these PCA cluster groups (Fig. 3C). The middle group displayed high heterozygosity while the two outer groups were primarily homozygous. The homozygous cluster with scores greater than 0.04 on PC1 was composed primarily of early run individuals (hereafter denoted as ‘early’ allele), while the homozygous cluster with scores less than -0.025 on PC1 was composed primarily of late run individuals (hereafter denoted as ‘late’ allele). Subvariation in the late allele, associated with the eleven odd year individuals that clustered together near PC1 = - 0.04, was mostly due to potential recombination in the 5` portion of *lrrc9* (Fig. S4).

Nearly all (95%) of the late allele homozygotes from the even lineage were late run fish. Similarly, the vast majority (91%) of early allele homozygotes from the even lineage were from the early run. Associations between run timing phenotypes and alleles were lower in the odd lineage. While all eleven odd lineage fish that were early allele homozygotes returned early, only a slight majority (62%) of the late allele homozygotes were from the late run. The heterozygote cluster was composed of a relatively equal proportion of early (44%) and late (56%) run fish across both lineages.

Overall patterns of LD within this region were similar for both lineages; however, the magnitude of linkage was consistently higher in the even lineage. The most prominent linkage block was located around the *lrrc9* gene (~72.70–72.74 Mb; Fig. 3D). Within this block, linkage was significantly higher in the even year ($\overline{r^2}$ = 0.62) than the odd year ($\overline{r^2}$ = 0.47; t = 31.72, p < 0.001, df=22,158). The linkage block corresponds closely with *lrrc9* but extends past the gene margins slightly further than the region of elevated F_{ST}. We documented two additional linkage blocks within this region. The first block, ~27 Kb in size, was located upstream of *lrrc9*. This block spanned part of the kif15 gene, an intergenic region, and the pseudogene *LOC124046521*. Linkage within this region was on average less than within the *lrrc9* linkage block for both lineages ($\bar{r}^2_{even}$ = 0.43, $\bar{r}^2_{odd}$= 0.37). The second block, a ~17.3 Kb linkage block just upstream of the *six6a* gene (Fig 3D) coincided with the elevated F_{ST} values upstream of

six6a in the even lineage (Fig. 3A). Linkage within this block was also significantly higher in the even year ($\overline{r^2}$= 0.51) than the odd year ($\overline{r^2}$ = 0.41; t = 4.25, p < 0.001, df=1157.3).

The *lrrc9* gene in pink salmon is 24.3 kb in size, has an aligned length of 5219 bp, cds length of 4,536 bp and a protein length of 1,511 amino acids. There are two transcript variants predicted for the gene (XM_046364572.1, XM_046364573.1) which differ by an 18 bp indel at position 321; however, their predicted protein sequence is identical (XP_046220528.1, XP_046220529.1). Within both lineages there were 110 SNPs identified in the *lrrc9* gene, 79 in the even lineage and 86 in the odd lineage. Of those SNPs, 11 SNPs were in the cds. There were 7 SNPs between the 'early' and 'late' alleles that were highly differentiated (F_{ST} > 0.63), none of which were located in cds or exons.

Comparing the genomic region containing lrrc9 in sockeye and pink salmon and comparing lrrc9 across species

We identified 544 SNPs in pink salmon and 536 SNPs in sockeye salmon within the genomic region containing *lrrc9* when both species were aligned to the chum salmon genome. Samples grouped strongly by species in a PCA (71% of variance explained by PC1), while clusters corresponding to homozygous ('early' and 'late' alleles) and heterozygous genotypes within each species were differentiated on PC2 (7.33% of variation; Fig. S5). Patterns of genetic differentiation between alleles differed between the two species, with sockeye salmon containing fixed or nearly fixed differences at the *lrrc9* gene and also in three other genes, *dhrs7*, *pcnx4*, and *six6a*, while fixed or nearly fixed differences in pink salmon were isolated to the *lrrc9* gene (Fig. 4A). The SNPs were distributed similarly between species within the *lrrc9* gene located between 28.13 and 28.15 Mb (388 SNPs in pink and 343 SNPs in sockeye), with many SNPs found in or near the exons (Fig. 4B). There was no evidence of conservation of highly differentiated SNPs between the species, as no SNPs in the *lrrc9* gene were fixed or nearly fixed between polymorphic loci that were shared between the species. However, we did document two such SNPs in initial alignments of sockeye and pink salmon to the pink salmon genome (data not shown).

The size of the *lrrc9* gene varies among salmon species. The gene in sockeye salmon is similar in size to pink salmon (22.7 kb), but in chum salmon is nearly twice the size (~41 kb). This difference in size and the number of indels interspersed through the gene likely impacts the efficacy of alignment of pink and sockeye salmon short read sequences to the *lrrc9* gene in the chum salmon reference genome. There are two protein isoforms predicted for the chum genome with substantial differences at the 3' end of the cds. The X1 isoform is most similar to the proteins predicted for pink and sockeye salmon and was used for comparisons. The lrrc9 protein is 1,511 amino acids long and showed no difference between pink and sockeye salmon. For pink salmon there were 5 SNPs with high differentiation (F_{ST} ≥ 0.8) between 'early' and 'late' alleles that occurred in the cds, one of which resulted in a non-synonomous substitution (Gly->Ser) at position 1,392. For sockeye there were three SNPs with high differentiation between run timing alleles within the cds, one causing a non-synonmous substitution (Met->Ile) at position 1,473.

Discussion:

The Lrrc9 gene region is associated with run timing in pink salmon

Our study shows a major effect locus influences run timing in pink salmon. The *lrrc9* region of chromosome 10 is strongly associated with run timing in the even-year lineage of Auke Creek pink salmon. Within this region, the strongest association is a ~40 kb stretch of chromosome 10 characterized by high LD. This region contains two discrete alleles, which we term “early” and “late” alleles, and variation at these alleles is largely conserved between the even and odd lineages. Notably, even though the association between run timing and *lrrc9* genotype is much stronger in the even-year lineage, all of the early allele homozygotes in the odd-year lineage had early run timing.

Pink salmon exhibit the highest straying rates among Pacific salmon species and therefore the lowest levels of natural population structure, with genetically similar populations spanning 100s of kilometers. However, our study illustrates that major adaptive genetic differences can occur on very small temporal scales in the absence of spatial differences in pink salmon. Specifically, early and late runs of pink salmon in Auke Creek are separated by ~20 days, spawn primarily in the same 50 m stream reach, and show little genetic divergence across most of their genomes. In contrast, the only other known major effect locus linked with run timing in salmon (*greb1l*) is found both in Chinook and in steelhead, and differs between conspecific runs that migrate into freshwater months apart and display significant neutral differentiation (F_{ST} ~0.03, Prince et al. 2017). Additionally, *greb1L*, is associated with large differences in reproductive maturation schedules and residence times in freshwater (Waples et al 2022).

The substantial differences in the degree of association between the *lrrc9* region and run timing in even and odd lineage pink salmon is intriguing given that each lineage contains both the early and late alleles. One potential explanation for this pattern is that genetic variants at loci distant from *lrrc9* also influence run timing, and do so differently in the two lineages. It is not uncommon for major effect loci to explain varying amounts of phenotypic variation in different populations (Narum et al. 2023). Although no other large effect loci in the odd lineage nor smaller peaks of association in the even lineage were observed, epistasis could affect run timing because different coadapted gene complexes have been shown to affect marine survival in the odd and even lineages in Auke Creek (Gharrett and Smoker 1991, 1999). Analyzing the genetic basis of run timing and *lrrc9* variation in other pink salmon populations from both lineages would shed light on this somewhat puzzling pattern, as would experimental crosses to place *lrrc9* genotypes on different genetic backgrounds.

The lrrc9 gene region is associated with run timing in sockeye salmon

The SNPs in the *lrrc9* gene region were highly divergent between sockeye salmon ecotypes (beach and tributary spawners). Divergence at the *lrrc9* region has been documented between beach and tributary spawning sockeye salmon throughout their range from the Columbia River through Canada, Alaska, and Russia (Veale and Russello 2017, Larson et al. 2019). However, it has been difficult to link this divergence to functional differences between the spawning types because they differ in many traits. The beach spawning ecotype is more deep-bodied, larger, and these fish generally spawn later than the tributary spawning ecotype (Quinn et al. 2001, Hilborn et al. 2003, Schindler et al. 2010). Previous genetic studies have also found that these spawning types display divergence in the genes of the major histocompatibility complex, which is involved in immune system function (Larson et al. 2014).

The conserved patterns of divergence at *lrrc9* in pink and sockeye salmon provide evidence that *lrrc9* is linked to variation in run timing in both species. Run timing in sockeye salmon is strongly associated with temperature (Schindler et al. 2013, Lisi et al. 2013) and is thought to be optimized to allow eggs to develop such that, on average, emigration timing of juveniles matches productivity in their rearing environment (Hodgson and Quinn 2002). Tributary-spawning sockeye salmon generally spawn earlier than beach-spawning sockeye salmon because tributaries are cooler on average, necessitating longer incubation of their eggs to fully develop (Hilborn et al. 2003). An exception to these patterns are sockeye salmon that spawn on island beaches in Lake Iliamna. These island beach sockeye are phenotypically similar to other beach spawners but spawn earlier, likely because embryos must hatch prior to lake freeze-up since they are not insulated by groundwater inputs like other beach spawners (Stewart et al. 2003). Notably, these early spawning island beach populations contain *lrrc9* genotypes that are similar to those in tributaries rather than those at other mainland beaches (Larson et al. 2019, Euclide et al. 2023, Veale and Russello 2017), providing further evidence that *lrrc9* is associated with run timing in sockeye salmon. The revelation that *lrrc9* is a major effect locus for run timing in sockeye and associated with spawning habitat contextualizes the underlying genetic variation within Lake Washington sockeye salmon after reintroduction and what might promote reproductive isolation within the system (Gustafson et al. 2001, Hendry et al. 2000).

Evolutionary history and function of the lrrc9 region

Comparisons among pink and sockeye salmon suggest that selection on the *lrrc9* gene to form putative early and late run alleles occurred independently. Alleles are more similar within species, and the species share very few polymorphisms. Additionally, patterns of variation are somewhat different across the genomic region. Pink salmon display a smaller area of elevated differentiation centered around the *lrrc9* gene while sockeye salmon display a wider region of elevated differentiation that includes the *dhrs7*, *pcnx4*, and *six6* genes in addition to *lrrc9*. This variation in sockeye salmon appears to be consistent across the range, as it has been documented from Alaska to southern Canada (Euclide et al. 2023, Tigano and Russello 2022). A molecular clock analysis suggested that the early and late alleles in sockeye salmon diverged approximately 4 million years ago (Veale and Russello 2017), well after the hypothesized speciation of sockeye and pink salmon (McPhail 1997). In pink salmon, these results suggest that the early and late alleles diverged prior to the split of even and odd lineages, as there are very few genetic differences in the alleles between the lineages. While it is unclear when pink salmon lineages diverged, it is very likely that divergence occurred prior to the last glacial maximum (> 18,000 years ago) but well after pink salmon diverged from other salmon species (Churikov and Gharrett 2002). The number of polymorphisms per base pair is similar between *lrrc9* alleles in sockeye and pink salmon, suggesting that alleles in pink salmon also diverged around the same time period (~4 MYA).

The apparent independent selection events at the *lrrc9* region in sockeye and pink salmon are similar to patterns documented for *greb1l* in Chinook salmon and steelhead trout, with alleles more similar within than between species (Prince et al. 2017). Additionally, *greb1l* variation displays a similar pattern where one species (Chinook salmon) contains a wider peak of high differentiation (Waples et al. 2022). The differences in the peak widths observed in both *lrrc9* and *greb1l*, likely provides additional evidence for independent selection events that occurred

on different standing genetic variation, and also that the overlapping peaks in candidate gene regions provide strong evidence that those regions are targets of selection.

Many recent studies have highlighted the importance of structural genomic variants such as inversions for facilitating adaptive divergence with gene flow (reviewed in Wellenreuther and Bernatchez 2018). However, the *lrrc9* alleles may not be associated with an inversion. Euclide et al. (2023) suggested that *lrrc9* alleles could represent an inversion based primarily on LD patterns, a relatively discrete break at the margins of the highly differentiated region, and genotype patterns in local PCAs. However, Tigano and Russello (2022) conducted long read sequencing on *lrrc9* alleles in sockeye salmon and did not find evidence of an inversion. Substructure observed in the odd-year lineage late allele PCA and allele heatmap may also reflect recombination within the lrrc9 gene. Additionally, the relatively small size of the highly differentiated regions and the fact that they differ in size in each species provides further evidence that the alleles are not underlain by an inversion. Similar lines of evidence have been used to hypothesize that *greb1l* alleles were not created by an inversion (Waples et al. 2022).

The small number of documented non-synonymous mutations between *lrrc9* alleles in pink and sockeye salmon, none of which were shared, suggests that functional changes associated with different alleles may be regulatory. Studies of *greb1l* have come to a similar conclusion, as there are very few non-synonymous mutations in the diverged region and the peak of association is found in an intergenic region (Waples et al. 2022).

Migration is a complex process involving many unrelated physiological pathways (Dingle 2014). It is logical to conclude that if only one major region of the genome diverged between run timing phenotypes, that region must be involved in regulating a large number of physiological processes. Recent research in Atlantic salmon has revealed that a major effect locus involved in age-at-maturity (*vgll3*) is a transcription cofactor that upregulates a number of genes controlling multiple processes such as androgen production and cellular energy (Verta et al. 2024). It is possible that variation at *lrrc9* similarly controls a large number of genes involved in run timing related traits. To test this, functional genomic studies are necessary.

Although the clearest signals of divergence in sockeye and pink salmon are found in and near the *lrrc9* gene, there are some other genes in this region that are diverged in one or both of the species. Most notably, we observed high differentiation near the *six6* gene in sockeye salmon and lower but still elevated differentiation near this gene in pink salmon (also documented for sockeye in Tigano and Russello 2022 and Euclide et al. 2023). The *six6* gene has been associated with variation in age-at-maturity in sockeye salmon, Atlantic salmon, and steelhead trout (Waters et al. 2021, Mobley et al. 2021). *Six6* encodes a homeobox protein, which are involved in regulation of many genes associated with anatomical development. Sockeye salmon also displayed elevated divergence at the *dhrs7* and *pcnx4* genes. The *dhrs7* gene encodes an enzyme and is potentially involved in eye size (Ramdas et al. 2010), while *pcnx4* encodes a membrane protein.

Conclusions

For many decades geneticists studying salmon postulated that complex life-history variation was controlled by many genes of small effect, but increasingly genomic studies provide

evidence that these traits can be controlled by large-effect loci that are sometimes shared among species. Our study provides evidence that the same large-effect locus can facilitate adaptive divergence on small scales in multiple species of salmon. Our study also demonstrates how a multispecies approach can help illuminate the genetic basis of phenotypic and life history differences that would otherwise remain obscured due to confounding ecological variation. One area of future research is investigating the genetic basis of run timing in additional populations of pink salmon to determine if there is a consistently stronger signal in the even lineage, or if signals vary among lineages and populations. It is possible that in some cases, other genetic variants are primarily responsible for influencing run timing in certain populations of pink salmon.

The conservation implications of this study suggest that if run timing in sockeye and pink salmon is largely controlled by a single locus, these species may have an increased ability to adapt more quickly to changing environmental conditions. Specifically, in the absence of epistasis, with a single identified major target of selection, populations may be able to more quickly shift their run timing in response to environmental changes than if the trait was controlled by many loci of small effect. Additionally, since populations of both sockeye and pink salmon appear to contain at least some heterozygotes and alternate homozygotes, and since the alleles are evolutionarily old and therefore likely widespread, it is unlikely that the genetic variants will disappear.

Acknowledgements
We thank the numerous graduate students, contractors, and NMFS staff who have provided foundational research and collected samples at the Auke Creek weir, especially John Joyce, Sidney Taylor, and Bill Smoker. We also thank Curry Cunningham, Nicolas Lou, Greg Owens, and Peter Sudmant for useful discussions about the paper. This work was also supported in part by the high-performance computing and data storage resources SEDNA operated by National Marine Fisheries Service.

References

Alexandrou MA, Swartz BA, Matzke NJ, Oakley TH. 2013 Genome duplication and multiple evolutionary origins of complex migratory behavior in salmonidae. *Molecular Phylogenetics and Evolution* 69, 514–523. (doi:https://doi.org/10.1016/j.ympev.2013.07.026)

Andrews S. 2010 FastQC: A quality control tool for high throughput sequence data. [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Beacham TD, McIntosh B, MacConnachie C, Spilsted B, White BA. 2012 Population structure of pink salmon (*Oncorhynchus gorbuscha*) in British Columbia and Washington, determined with microsatellites. *Fisheries Bulletin* 110, 242–256.

Beacham T, Murray C. 1990 Temperature, egg size, and development of embryos and alevins of five species of pacific salmon: A comparative analysis. *Transactions of The American Fisheries Society* 119, 927–945. (doi:10.1577/1548-8659(1990)119<0927:TESADO>2.3.CO;2)

Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. (doi:10.1093/bioinformatics/btu170)

Campbell S, Butler V. 2010 Archaeological evidence for resilience of Pacific Northwest salmon populations and the socioecological system over the last 7,500 years. *Ecology and Society* 15, 17. (doi:10.5751/ES-03151-150117)

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. *Molecular Ecology Resources* 15, 855–867. (doi:https://doi.org/10.1111/1755-0998.12357)

Carlson SM, Seamons TR. 2008 A review of quantitative genetic components of fitness in salmonids: Implications for adaptation to future change. *Evolutionary Applications* 1, 222–238. (doi:10.1111/j.1752-4571.2008.00025.x)

Cauwelier E, Gilbey J, Sampayo J, Stradmeyer L, Middlemas SJ. 2018 Identification of a single genomic region associated with seasonal river return timing in adult scottish atlantic salmon (*Salmo salar*), using a genome-wide association study. *Canadian Journal of Fisheries and Aquatic Sciences* 75, 1427–1435. (doi:10.1139/cjfas-2017-0293)

Chen S, Zhou Y, Chen Y, Gu J. 2018 fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. (doi:10.1093/bioinformatics/bty560)

Christensen KA *et al.* 2021 The pink salmon genome: Uncovering the genomic consequences of a two-year life cycle. *PLoS One* 16, e0255752. (doi:10.1371/journal.pone.0255752)

Churikov D, Gharrett AJ. 2002 Comparative phylogeography of the two pink salmon broodlines: An analysis based on a mitochondrial DNA genealogy. *Molecular Ecology* 11, 1077–1101. (doi:https://doi.org/10.1046/j.1365-294X.2002.01506.x)

Dingle H. 2014 Physiology of migration. In *Migration: The biology of life on the move*, Oxford University Press. (doi:10.1093/acprof:oso/9780199640386.003.0006)

Dingle H, Drake VA. 2007 What is migration? *BioScience* 57, 113–121. (doi:10.1641/B570206)

Echave JD, Manhard CV, Smoker WW, Adkison MD, Gharrett AJ. 2017 Out crosses between seasonally different segments of a pacific salmon population reveal local adaptation. *Environmental Biology of Fishes* 100, 1469–1481. (doi:10.1007/s10641-017-0657-3)

Eliason EJ *et al.* 2011 Differences in thermal tolerance among sockeye salmon populations. *Science* 332, 109–112. (doi:10.1126/science.1199158)

Euclide PT, Larson WA, Shi Y, Gruenthal K, Christensen KA, Seeb J, Seeb L. 2023 Conserved islands of divergence associated with adaptive variation in sockeye salmon are maintained by multiple mechanisms. *Molecular Ecology* 00, 1–21. (doi:10.1111/mec.17126)

Ewels P, Magnusson M, Lundin S, Käller M. 2016 MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32, 3047–3048. (doi:10.1093/bioinformatics/btw354)

Fox EA, Wright AE, Fumagalli M, Vieira FG. 2019 ngsLD: evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics* 35, 3855–3856. (doi:10.1093/bioinformatics/btz200)

Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB. 2011 Extent and scale of local adaptation in salmonid fishes: Review and meta-analysis. *Heredity* 106, 404–420. (doi:10.1038/hdy.2010.167)

Fukushima M, Quinn TJ, Smoker WW. 1998 Estimation of eggs lost from superimposed pink salmon (*Oncorhynchus gorbuscha*) redds. *Canadian Journal of Fisheries and Aquatic Sciences* 55, 618–625. (doi:10.1139/f97-260)

Fumagalli M, Vieira FG, Linderoth T, Nielsen R. 2014 ngsTools: Methods for population genetics analyses from next-generation sequencing data. *Bioinformatics* 30, 1486–1487. (doi:10.1093/bioinformatics/btu041)

Garcia de Leaniz C *et al.* 2007 A critical review of adaptive genetic variation in Atlantic salmon: Implications for conservation. *Biological Reviews* 82, 173–211. (doi:10.1111/j.1469-185X.2006.00004.x)

Gharrett A, Smoker W, Reisenbichler R, Taylor S. 1999 Outbreeding depression in hybrids between odd- and even-year pink salmon. *Aquaculture* 173, 117–129. (doi:10.1016/S0044-8486(98)00480-3)

Gharrett AJ, Lane S, McGregor AJ, Taylor SG. 2001 Use of a genetic marker to examine genetic interaction among subpopulations of pink salmon (*Oncorhynchus gorbuscha*). *Genetica* 111, 259–267. (doi:10.1023/A:1013791314900)

Gharrett AJ, Joyce J, Smoker WW. 2013 Fine-scale temporal adaptation within a salmonid population: Mechanism and consequences. *Molecular Ecology* 22, 4457–4469. (doi:10.1111/mec.12400)

Gilk SE, Wang IA, Hoover CL, Smoker WW, Taylor SG, Gray AK, Gharrett AJ. 2004 Outbreeding depression in hybrids between spatially separated pink salmon, *Oncorhynchus gorbuscha*, populations: Marine survival, homing ability, and variability in family size. *Environmental Biology of Fishes* 69, 287–297. (doi:10.1023/B:EBFI.0000022888.28218.c1)

Gu Z. 2022 Complex heatmap visualization. *iMeta* 1, e43. (doi:10.1002/imt2.43)

Gustafson R, Waples R, Kalinowski S, Winans G. 2001. Evolution of Sockeye Salmon Ecotypes. *Science* 291,251-252(2001).(doi:10.1126/science.291.5502.251B)

Hawkins, S.L., N.V. Varnavskaya, E.A. Matzak, V.V. Efremov, C.M. Guthrie III, R.L. Wilmot, H. Mayama, F. Yamazaki, and A.J. Gharrett. 2002. Contrasting population structures of Asian odd- and even-broodline pink salmon. *Journal of Fish Biology* 60,370-388.. (doi:10.1111/j.1095-8649.2002.tb00287.x)

Halffman CM, Potter BA, McKinney HJ, Finney BP, Rodrigues AT, Yang DY, Kemp BM. 2015 Early human use of anadromous salmon in North America at 11,500 y ago. *Proceedings of the National Academy of Sciences* 112, 12344–12348. (doi:10.1073/pnas.1509747112)

Heard WR. 1991 Pacific salmon life histories. In (eds C Groot, L Margolis), pp. 119–230. University of British Columbia Press.

Hebert KP, Goddard PL, Smoker WW, Gharrett AJ. 1998 Quantitative genetic variation and genotype by environment interaction of embryo development rate in pink salmon (*Oncorhynchus gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences* 55, 2048–2057. (doi:10.1139/f98-084)

Hendry A, Wenburg J, Bentzen P, Volk E, Quinn T. 2000. Rapid evolution of reproductive isolation in the wild: Evidence from introduced salmon.*Science* 290,516-518. (do:10.1126/science.290.5491.516)

Hess JE, Ackerman MW, Fryer JK, Hasselman DJ, Steele CA, Stephenson JJ, Whiteaker JM, Narum SR. 2016 Differential adult migration-timing and stock-specific abundance of steelhead in mixed stock assemblages. *ICES Journal of Marine Science* 73, 2606–2615. (doi:10.1093/icesjms/fsw138)

Hilborn R, Quinn TP, Schindler DE, Rogers DE. 2003 Biocomplexity and fisheries sustainability. *Proceedings of the National Academy of Sciences* 100, 6564–6568. (doi:10.1073/pnas.1037274100)

Hilderbrand G, Schwartz C, Robbins CT, Jacoby ME, Hanley TA, Arthur S, Servheen C. 1999 The importance of meat, particularly salmon, to body size, population productivity, and conservation of North American brown bears. *Canadian Journal of Zoology* 77, 132–138. (doi:10.1139/cjz-77-1-132)

Hodgson S, Quinn TP. 2002 The timing of adult sockeye salmon migration into fresh water: Adaptations by populations to prevailing thermal regimes. *Canadian Journal of Zoology* 80, 542–555. (doi:10.1139/z02-030)

Horn RL, Narum SR. 2023 Genomic variation across Chinook salmon populations reveals effects of a duplication on migration alleles and supports fine scale structure. *Molecular Ecology* 32, 2818–2834. (doi:10.1111/mec.16895)

Jun G, Wing MK, Abecasis GR, Kang HM. 2015 An efficient and scalable analysis framework for variant extraction and refinement from population scale DNA sequence data. *Genome Research* (doi:10.1101/gr.176552.114)

Kitano T, Matsuoka N, Saitou N. 1997 Phylogenetic relationship of the genus *Oncorhynchus* species inferred from nuclear and mitochondrial markers. *Genes and Genetic Systems* 72, 25–34. (doi:10.1266/ggs.72.25)

Koch IJ, Narum SR. 2021 An evaluation of the potential factors affecting lifetime reproductive success in salmonids. *Evolutionary Applications* 14, 1929–1957. (doi:10.1111/eva.13263)

Korneliussen TS, Albrechtsen A, Nielsen R. 2014 ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics* 15, 356. (doi:10.1186/s12859-014-0356-4)

Kovach RP, Gharrett AJ, Tallmon DA. 2012 Genetic change for earlier migration timing in a pink salmon population. *Proceedings of the Royal Society B: Biological Sciences* 279, 3870–3878. (doi:10.1098/rspb.2012.1158)

Kovach JEAE Ryan P. AND Joyce. 2013 Earlier migration timing, decreasing phenotypic variation, and biocomplexity in multiple salmonid species. *PLOS ONE* 8, 1–10. (doi:10.1371/journal.pone.0053807)

Kovach RP, Ellison SC, Pyare S, Tallmon DA. 2015 Temporal patterns in adult salmon migration timing across southeast alaska. *Global Change Biology* 21, 1821–1833. (doi:10.1111/gcb.12829)

Larson WA, Seeb JE, Dann TH, Schindler DE, Seeb LW. 2014 Signals of heterogeneous selection at an MHC locus in geographically proximate ecotypes of sockeye salmon. *Molecular Ecology* 23, 5448–5461. (doi:10.1111/mec.12949)

Larson WA, Dann TH, Limborg MT, McKinney GJ, Seeb JE, Seeb LW. 2019 Parallel signatures of selection at genomic islands of divergence and the major histocompatibility complex in ecotypes of sockeye salmon across alaska. *Molecular Ecology* 28, 2254–2271. (doi:10.1111/mec.15082)

Levi T, Wheat RE, Allen JM, Wilmers CC. 2015 Differential use of salmon by vertebrate consumers: Implications for conservation. *PeerJ* 3, e1157. (doi:10.7717/peerj.1157)

Li H *et al.* 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. (doi:10.1093/bioinformatics/btp352)

Li H, Durbin R. 2009 Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics* 25, 1754–1760. (doi:10.1093/bioinformatics/btp324)

Limborg MT, Waples RK, Seeb JE, Seeb LW. 2014 Temporally isolated lineages of pink salmon reveal unique signatures of selection on distinct pools of standing genetic variation. *Journal of Heredity* 105, 741–751. (doi:10.1093/jhered/esu063)

Manhard CV, Joyce JE, Gharrett AJ. 2017 Evolution of phenology in a salmonid population: A potential adaptive response to climate change. *Canadian Journal of Fisheries and Aquatic Sciences* 74, 1519–1527. (doi:10.1139/cjfas-2017-0028)

Manhard CV, Joyce JE, Smoker WW, Gharrett AJ. 2017 Ecological factors influencing lifetime productivity of pink salmon (*Oncorhynchus gorbuscha*) in an Alaskan stream. *Canadian Journal of Fisheries and Aquatic Sciences* 74, 1325–1336. (doi:10.1139/cjfas-2016-0335)

Manhard CV, Adkison MD, Hard JJ, Smoker WW, Gharrett AJ. 2018 Local adaptation of phenology revealed in outcrosses between spawning segments of a salmonid population. *Molecular Ecology* 27, 4698–4710. (doi:10.1111/mec.14908)

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

McPhail JD. 1997 The origin and speciation of *Oncorhynchus* revisited. In *Pacific salmon & their ecosystems: Status and future options* (eds DJ Stouder, PA Bisson, RJ Naiman), pp. 29–38. Boston, MA: Springer US. (doi:10.1007/978-1-4615-6375-4_4)

Meisner J, Albrechtsen A. 2018 Inferring population structure and admixture proportions in low-depth NGS data. *Genetics* 210, 719–731. (doi:10.1534/genetics.118.301336)

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

Mobley KB *et al.* 2021 Maturation in Atlantic salmon (*Salmo salar, Salmonidae*): A synthesis of ecological, genetic, and molecular processes. *Reviews in Fish Biology and Fisheries* 31, 523–571. (doi:10.1007/s11160-021-09656-w)

Mortensen DG, Wertheimer AC, Maselko JM, Taylor SG. 2002 Survival and straying of Auke Creek, Alaska, pink salmon marked with coded wire tags and thermally induced otolith marks. *Transactions of the American Fisheries Society* 131, 14–26. (doi:10.1577/1548-8659(2002)131<0014:SASOAC>2.0.CO;2)

Narum, S. R., Horn, R., Willis, S., Koch, I., & Hess, J. (2024). Genetic variation associated with adult migration timing in lineages of Steelhead and Chinook salmon in the Columbia River. *Evolutionary Applications*, 17, e13626. (doi:10.1111/eva.13626)

Nichols KM, Kozfkay CC, Narum SR. 2016 Genomic signatures among *Oncorhynchus nerka* ecotypes to inform conservation and management of endangered sockeye salmon. *Evolutionary Applications* 9, 1285–1300. (doi:10.1111/eva.12412)

Oke KB, Cunningham CJ, Quinn TP, Hendry AP. 2019 Independent lineages in a common environment: The roles of determinism and contingency in shaping the migration timing of even- versus odd-year pink salmon over broad spatial and temporal scales. *Ecology Letters* 22, 1547–1556. (doi:10.1111/ele.13337)

O'Malley KG, Ford MJ, Hard JJ. 2010 Clock polymorphism in pacific salmon: Evidence for variable selection along a latitudinal gradient. *Proceedings of the Royal Society B: Biological Sciences* 277, 3703–3714. (doi:10.1098/rspb.2010.0762)

Pearse DE *et al.* 2019 Sex-dependent dominance maintains migration supergene in rainbow trout. *Nature Ecology & Evolution* 3, 1731–1742. (doi:10.1038/s41559-019-1044-6)

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. *Science Advances* 3, e1603198. (doi:10.1126/sciadv.1603198)

Quinn TP, Wetzel L, Bishop S, Overberg K, Rogers DE. 2001 Influence of breeding habitat on bear predation and age at maturity and sexual dimorphism of sockeye salmon populations. *Canadian Journal of Zoology* 79, 1782–1793. (doi:10.1139/z01-134)

Ramdas LMEAI Wishal D. AND van Koolwijk. 2010 A genome-wide association study of optic disc parameters. *PLOS Genetics* 6, 1–12. (doi:10.1371/journal.pgen.1000978)

Ricker WE. 1972 The stock concept in pacific salmon. H. R. MacMillian lectures in fisheries. In (eds RC Simon, PA Larakin), pp. 16–160. Vancouver: University of British Columbia.

Schindler DE, Armstrong JB, Bentley KT, Jankowski K, Lisi PJ, Payne LX. 2013 Riding the crimson tide: Mobile terrestrial consumers track phenological variation in spawning of an anadromous fish. *Biology Letters* 9, 20130048. (doi:10.1098/rsbl.2013.0048)

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. (doi:10.1038/nature09060)

Schindler D, Scheuerell M, Moore J, Gende S, Francis T, Palen W. 2003 Pacific salmon and the ecology of coastal ecosystems. *Frontiers in Ecology and the Environment* 1, 31–37. (doi:10.1890/1540-9295(2003)001[0031:PSATEO]2.0.CO;2)

Seeb LW, Waples RK, Limborg MT, Warheit KI, Pascal CE, Seeb JE. 2014 Parallel signatures of selection in temporally isolated lineages of pink salmon. *Molecular Ecology* 23, 2473–2485. (doi:10.1111/mec.12769)

Smoker WW, Gharrett AJ, Stekoll MS. 1998 Genetic variation of return date in a population of pink salmon: A consequence of fluctuating environment and dispersive selection. *Alaska Fishery Research Bulletin* 5, 46–54.

Stewart, I.J., Hilborn, R. and Quinn, T.P. 2003. Coherence of observed adult sockeye salmon abundance within and among spawning habitats in the Kvichak River watershed. *Alaska Fishery Research Bulletin* 10:28-41.

Tarpey CM, Seeb JE, McKinney GJ, Templin WD, Bugaev A, Sato S, Seeb LW. 2018 Single-nucleotide polymorphism data describe contemporary population structure and diversity in allochronic lineages of pink salmon (*Oncorhynchus gorbuscha*). *Canadian Journal of Fisheries and Aquatic Sciences* 75, 987–997. (doi:10.1139/cjfas-2017-0023)

Taylor SG. 2008 Climate warming causes phenological shift in pink salmon, *Oncorhynchus gorbuscha*, behavior at Auke Creek, Alaska. *Global Change Biology* 14, 229–235. (doi:10.1111/j.1365-2486.2007.01494.x)

Therkildsen NO, Palumbi SR. 2017 Practical low-coverage genomewide sequencing of hundreds of individually barcoded samples for population and evolutionary genomics in

nonmodel species. *Molecular Ecology Resources* 17, 194–208. (doi:10.1111/1755-0998.12593)

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. (doi:10.1126/science.aba9059)

Tigano A, Russello MA. 2022 The genomic basis of reproductive and migratory behaviour in a polymorphic salmonid. *Molecular Ecology* 31, 6588–6604. (doi:0.1111/mec.16724)

Linhoff M, Harton J, Cressman D, Martin B, Pan-Yun Ting J. 2001 Two distinct domains within CIITA mediate self-association: Involvement of the GTP-binding and leucine-rich repeat domains. *Molecular and Cellular Biology* 21, 3001–3011. (doi:10.1128/MCB.21.9.3001-3011.2001)

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012 Primer3—new capabilities and interfaces. *Nucleic Acids Research* 40, e115–e115. (doi:10.1093/nar/gks596)

Veale AJ, Russello MA. 2017 An ancient selective sweep linked to reproductive life history evolution in sockeye salmon. *Scientific Reports* 7, 1747. (doi:10.1038/s41598-017-01890-2)

Verta J-P, Moustakas-Verho J, Donner I, Frapin M, Ruokolainen A, Debes P, Erkinaro J, Primmer J. 2024 A complex mechanism translating variation of a simple genetic architecture into alternative life-histories. *bioRxiv* (doi: 10.1101/2024.01.05.574286).

Waples RS *et al.* 2022 Implications of large-effect loci for conservation: A review and case study with pacific salmon. *Journal of Heredity* 113, 121–144. (doi:10.1093/jhered/esab069)

Waters CD, Clemento A, Aykanat T, Garza JC, Naish KA, Narum S, Primmer CR. 2021 Heterogeneous genetic basis of age at maturity in salmonid fishes. *Molecular Ecology* 30, 1435–1456. (doi:10.1111/mec.15822)

Wellenreuther M, Bernatchez L. 2018 Eco-evolutionary genomics of chromosomal inversions. *Trends in Ecology & Evolution* 33, 427–440. (doi:10.1016/j.tree.2018.04.002)

Wertheimer AC, Heintz RA, Thedinga JF, Maselko JM, Rice SD. 2000 Straying of adult pink salmon from their natal stream following embryonic exposure to weathered exxon valdez crude oil. *Transactions of the American Fisheries Society* 129, 989–1004. (doi:10.1577/1548-8659(2000)129<0989:SOAPSF>2.3.CO;2)

Wickham H. 2016 Ggplot2: Elegant graphics for data analysis. 2nd edn. Cham, Switzerland: Springer International Publishing.

Willson MF, Halupka KC. 1995 Anadromous fish as keystone species in vertebrate communities. *Conservation Biology* 9, 489–497.

Tables:

Table 1: Collection information for the 288 male and female pink salmon from the early and late portions of the even and odd lineage return to Auke Creek used for whole genome sequencing.

Lineage	Run Timing	Year	Sex	Return Date	Number of Samples
Even	Early	2006	Female	3,4,5-Aug-2006	37
Even	Early	2006	Male	3,4,5-Aug-2006	37
Even	Late	2002	Male	6-Sep-2002	20
Even	Late	2002	Male	9-Sep-2002	22
Even	Late	2006	Female	6,7,9-Sep-2006	37
Odd	Early	2005	Female	22-Aug-2005	36
Odd	Early	2005	Male	22-Aug-2005	32
Odd	Late	2001	Male	11-Sept-2001	31
Odd	Late	2005	Female	7-Sep-2005	36

Table 2: Genes within the region of highest association with run timing on chromosome 10 of the pink salmon genome.

Gene	Description	Size (nt)	Start	End	Function
kif15	kinesin family member 15	15468	72594558	72610025	ATP binding, microtubule binding, micotubule motor activity
LOC124046512	ATP-dependent RNA helicase TDRD9-like	3915	72627791	72631705	ATP binding, RNA binding
rtn1a	reticulon 1a	50472	72659990	72710461	brain devo, endoplasmic reticulum tubular network formation, neuro differentiation
lrrc9	leucine rich repeat containing 9	24316	72710811	72735126	protenin binding
LOC124046514	pecanex-like protein 4	3906	72765087	72768992	
dhrs7	dehydrogenase/reductase (SDR family) member 7	7302	72773886	72781187	oxidoreductase activity
ppm1aa	protein phosphatase, Mg2+/Mn2+ dependent, 1Aa	19010	72781554	72800563	magnesium/manganese ion binding, protein serine/theronine phosphatase activity
six6a	SIX homeobox 6a	6122	72865091	72871212	DNA-binding transcription factor activity (RNA polymerase II-specific), RNA polymerase II cis-regulatory region sequence-specific DNA binding
mnat1	MNAT1 component of CDK activating kinase	90349	72885946	72976294	cyclin-dependent protein serine/theronine kinase activator activity

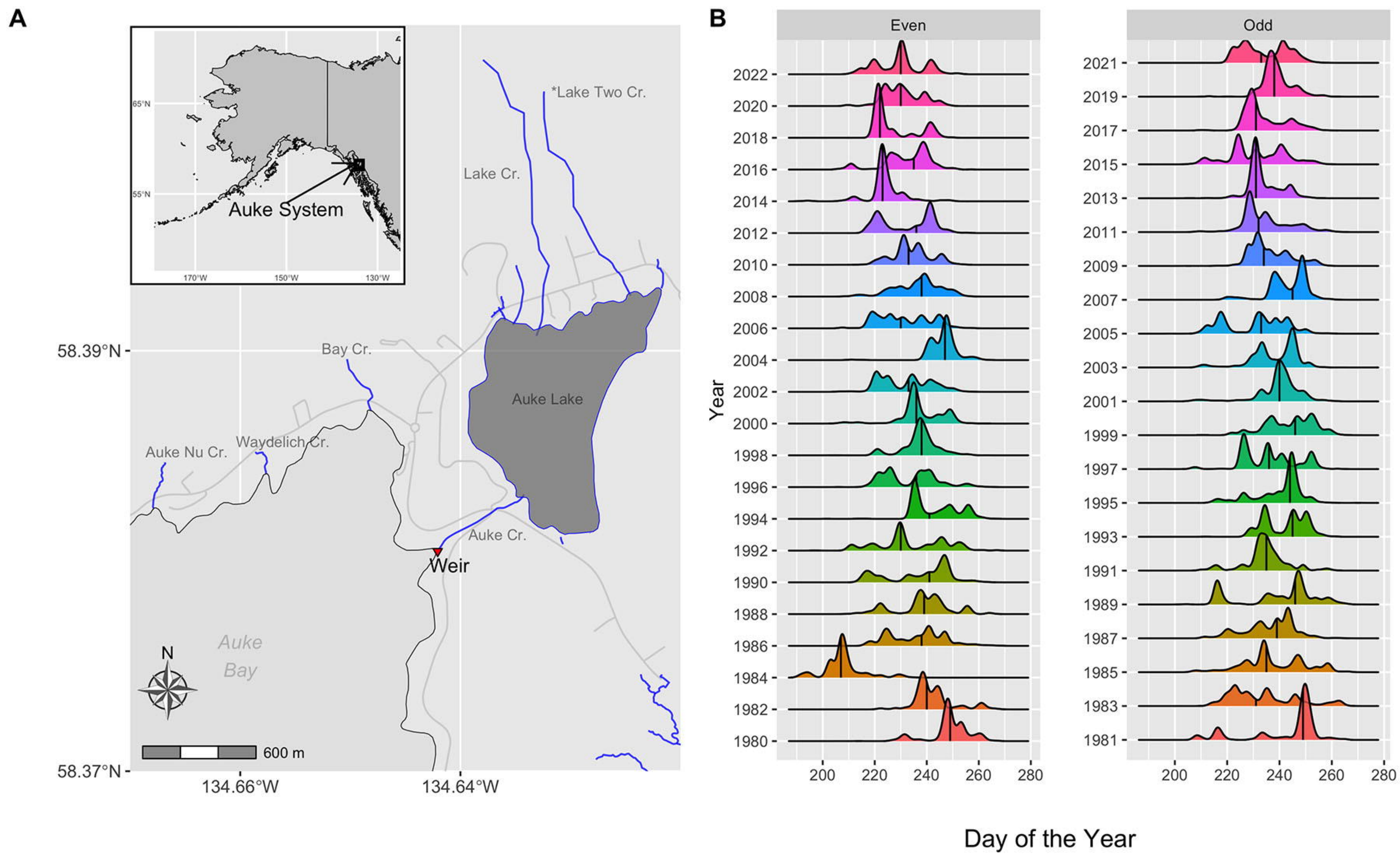
Figure legends

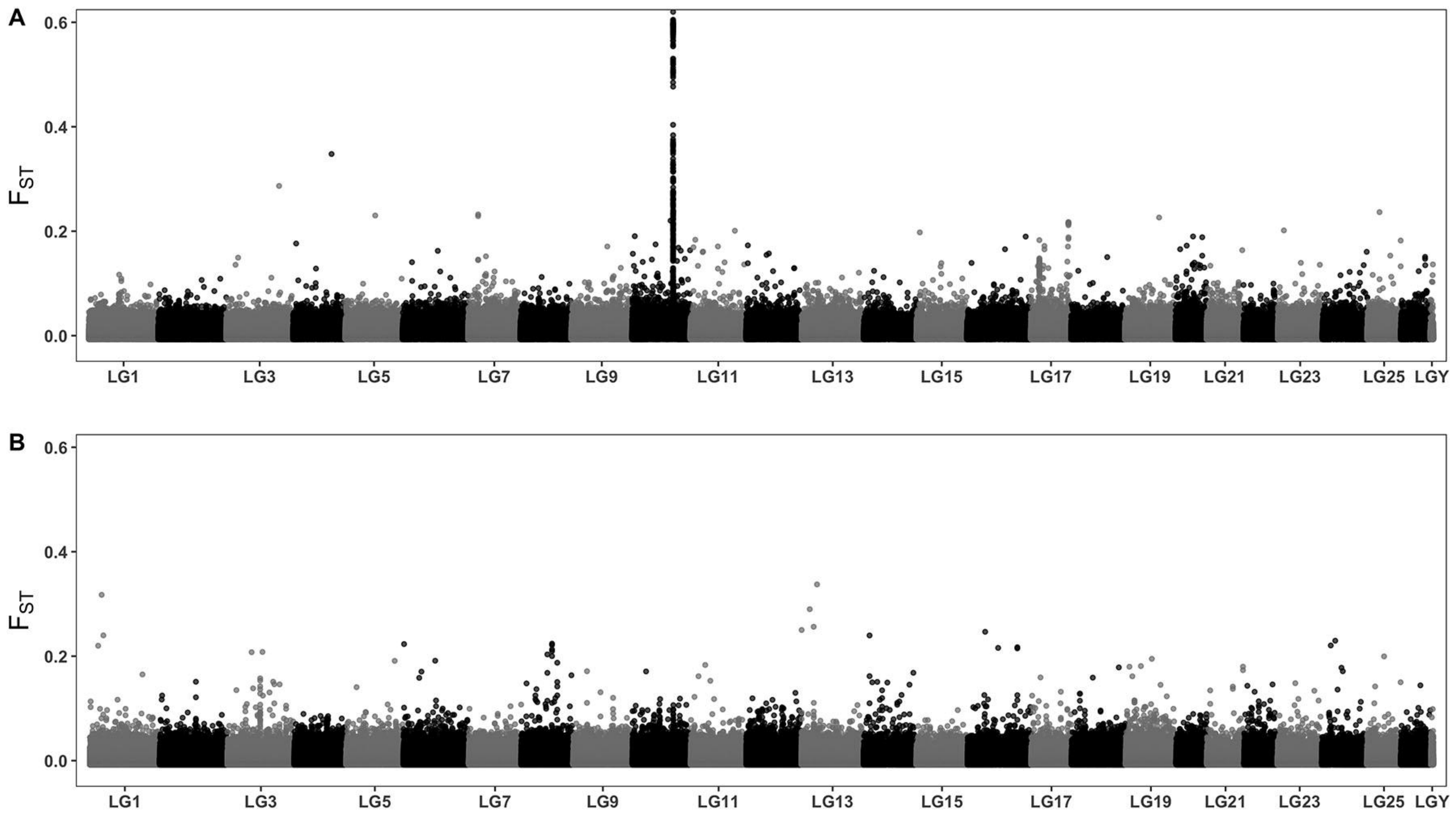
Figure 1: Return timing dynamics of Auke creek pink salmon. (A) Map of the Auke study system in Southeast Alaska. The majority of pink salmon spawn in Auke Creek, while some spawning occurs in the inlet tributary Lake Creek. All immigrant and emigrant pink salmon have been counted daily at the Auke Creek weir since 1980. (B) Density of return timing of adult pink salmon to the weir, with median return date indicated with vertical line.

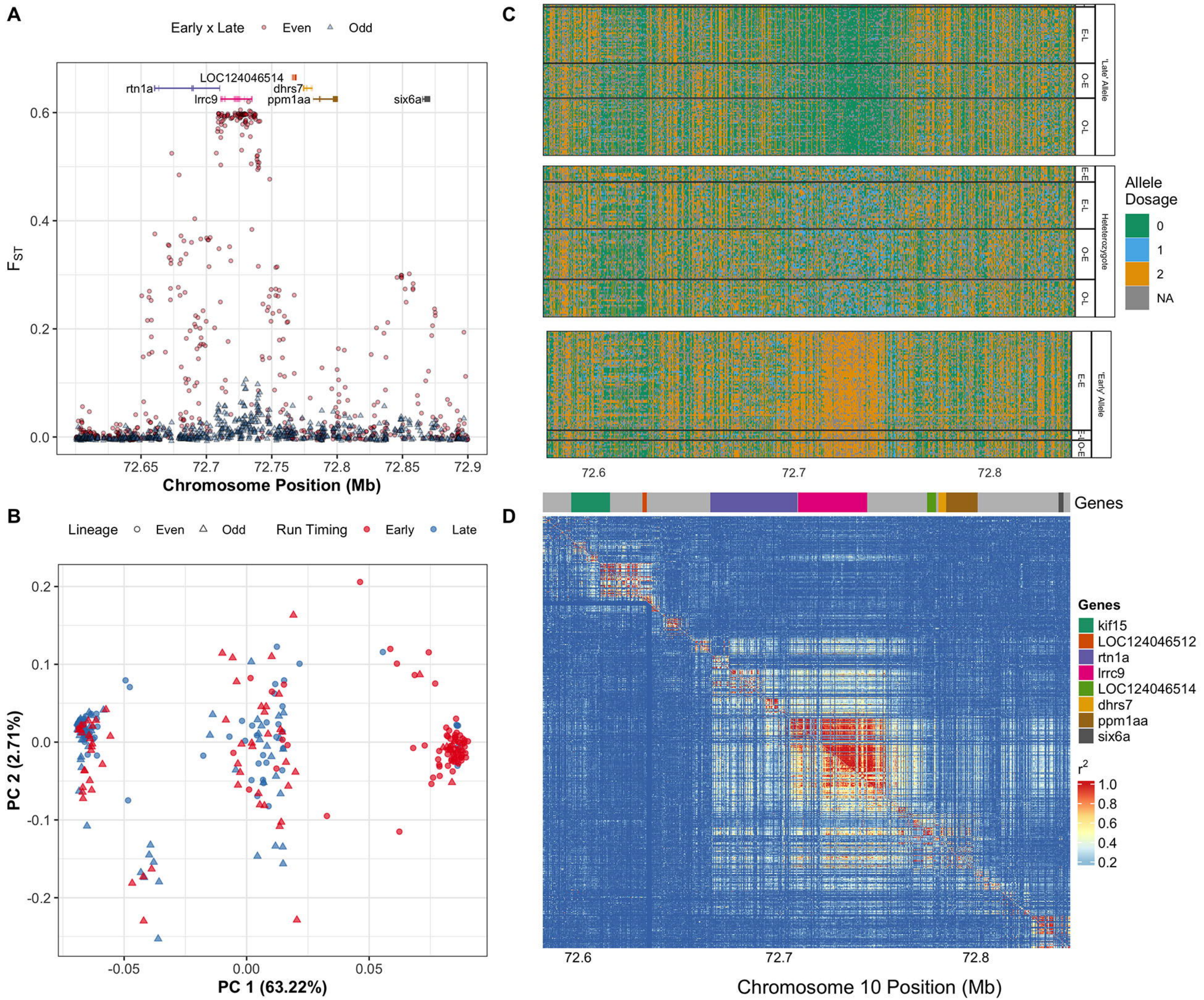
Figure 2: Manhattan plot of F_{ST} between early and late run pink salmon of the (A) even and (B) odd lineages in Auke Creek, Alaska. High divergence was observed for the even lineage on chromosome 10.

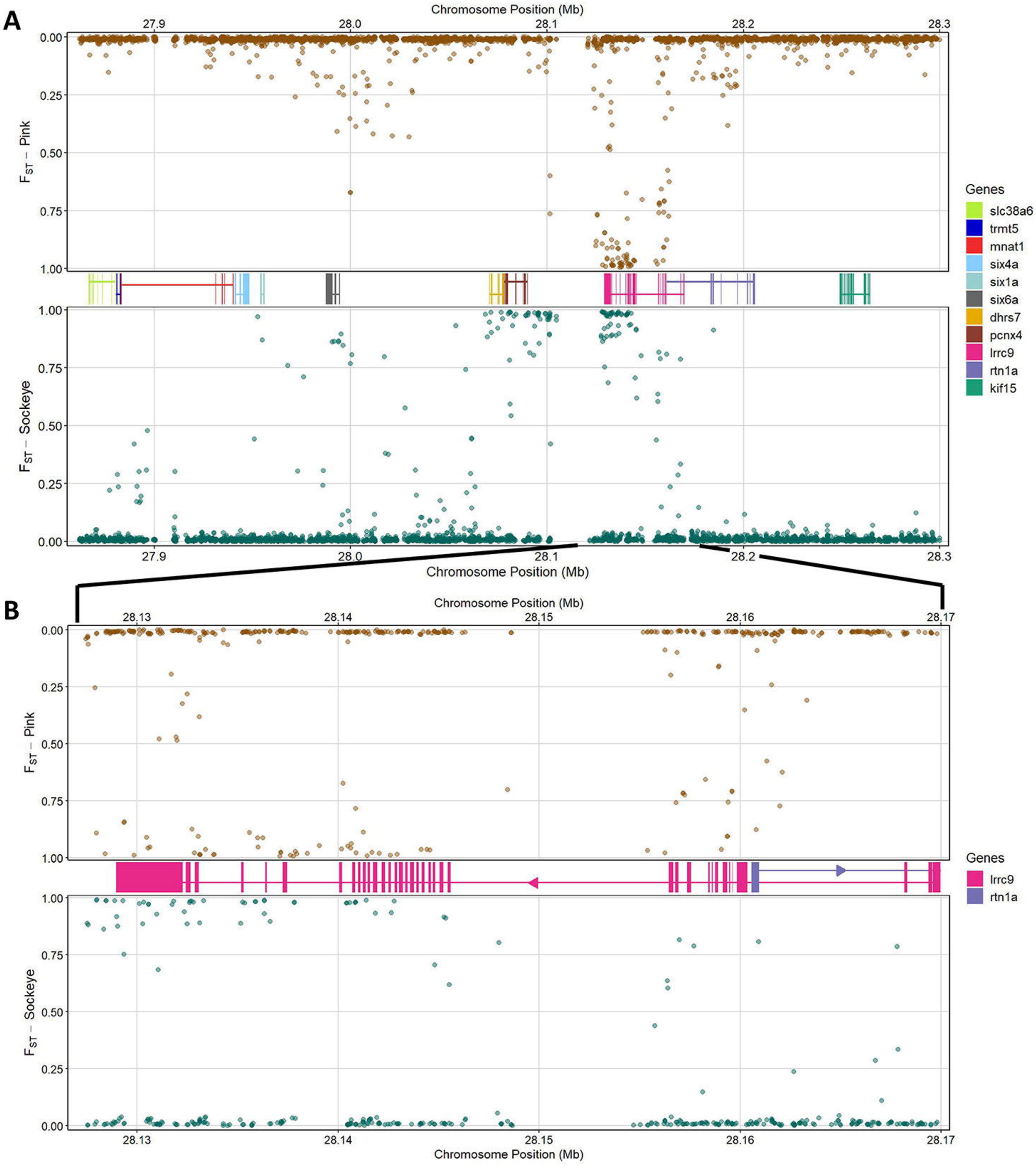
Figure 3: Chromosome 10 differentiation between early and late migrating pink salmon of the even and odd lineages in Auke Creek, Alaska. (A) Manhattan plot of pairwise F_{ST} between early and late run pink salmon of the even (red circles) and odd (blue triangles) lineages. (B) Principal component analysis including early even, late even, early odd, and late odd pink salmon. The three clusters represent the 'late' allele homozygotes (negative PC1 space), heterozygous (clustered at the origin), and 'early' allele homozygotes (positive PC1 space) groups depicted in the allele dosage plot (C). (C) SNP heatmap for pink salmon grouped by zygosity, lineage, and run timing. Each subplot represents one of the three clusters identified in (B) and within each subplot individuals are arranged by Even-Early (E-E), Even-Late (E-L), Odd-Early (O-E) and Odd-Late (O-L). There are only two E-E individuals with the late allele, and no O-L individuals with the early allele. (D) Linkage disequilibrium (r^2) in the even (upper triangle) and odd (lower triangle) lineages show multiple blocks, the largest of which occurs within the *lrrc9* gene.

Figure 4: Manhattan plots of pairwise F_{ST} between the 'early' and 'late' allele for pink (top plots, inverted F_{ST}) and sockeye (bottom plot) aligned to chromosome 35 of the chum salmon genome. A) Differentiation within 27.8 - 28.3 Mb. Genes within the region and the corresponding exons are denoted between plots. B) Differentiation within *lrrc9*. Both pink and sockeye salmon have many SNPs with F_{ST} ~ 1, whereas only sockeye have SNPs with F_{ST} ~ 1 within the *dhrs7* and *pcnx4* region. Both regions have elevated F_{ST} around six6a, but there is greater differentiation in sockeye.


A
65°N
55°N
Auke System
170°W
150°W
130°W
*Lake Two Cr.
Lake Cr.
Bay Cr.
Auke Lake
Auke Nu Cr.
Waydelich Cr.
Auke Cr.
Weir
Auke
Bay
N
600 m
58.39°N
58.37°N
134.66°W
134.64°W
B
Even
Odd
Year
2022
2020
2018
2016
2014
2012
2010
2008
2006
2004
2002
2000
1998
1996
1994
1992
1990
1988
1986
1984
1982
1980
2021
2019
2017
2015
2013
2011
2009
2007
2005
2003
2001
1999
1997
1995
1993
1991
1989
1987
1985
1983
1981
200
220
240
260
280
Day of the Year


A
B
F_{ST}
0.0
0.2
0.4
0.6
LG1
LG3
LG5
LG7
LG9
LG11
LG13
LG15
LG17
LG19
LG21
LG23
LG25
LGY


A
Early x Late
Even
Odd
LOC124046514
rtn1a
dhrs7
lrrc9
ppm1aa
six6a
F_{ST}
0.6
0.4
0.2
0.0
72.65
72.7
72.75
72.8
72.85
72.9
Chromosome Position (Mb)
B
Lineage
Even
Odd
Run Timing
Early
Late
PC 2 (2.71%)
0.2
0.1
0.0
-0.1
-0.2
-0.05
0.00
0.05
PC 1 (63.22%)
C
E-L
O-E
O-L
'Late' Allele
E-E
E-L
O-E
O-L
Heteterozygote
E-E
E-L
O-E
'Early' Allele
Allele Dosage
0
1
2
NA
72.6
72.7
72.8
Genes
D
Genes
kif15
LOC124046512
rtn1a
lrrc9
LOC124046514
dhrs7
ppm1aa
six6a
r^2
1.0
0.8
0.6
0.4
0.2
72.6
72.7
72.8
Chromosome 10 Position (Mb)


A
Chromosome Position (Mb)
27.9
28.0
28.1
28.2
28.3
0.00
0.25
0.50
0.75
1.00
F_{ST} – Pink
F_{ST} – Sockeye
Genes
slc38a6
trmt5
mnat1
six4a
six1a
six6a
dhrs7
pcnx4
lrrc9
rtn1a
kif15
B
28.13
28.14
28.15
28.16
28.17
Genes
lrrc9
rtn1a