GENOMIC EVIDENCE FOR HATHERY-INDUCED DOMESTICATION SELECTION IN CHINOOK SALMON, *ONCORHYNCHUS TSHAWYTSCHA*

by

NATASHA HOWE

Bachelor of Science, 2017
Seattle University
Seattle, WA

Submitted to the Graduate Faculty of the
College of Science and Engineering
Texas Christian University
In partial fulfillment of the requirements
for the degree of

Master of Science

May 2023

GENOMIC EVIDENCE FOR HATCHERY-INDUCED DOMESTICATION SELECTION IN CHINOOK SALMON, *ONCORHYNCHUS TSHAWYTSCHA*

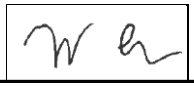
by

Natasha Howe

Dissertation approved:

Major Professor, Matthew Hale PhD

Charles D. Waters, PhD



Wesley Larson, PhD

For The College of Science and Engineering Provide the exact number of lines for your committee and the college signature. (e.g., committee = 4 college = 1 5 lines)

ACKNOWLEDGEMENTS

I am grateful to my major advisor, Dr. Matt Hale, for bringing me on with very little genetics knowledge, for teaching me basic and complex concepts over the past two years, and for reminding me of the importance of my own well-being. I am also extremely appreciative to my committee members, Dr. Wes Larson and Dr. Charlie Waters, for contributing their time and expertise both to this project and to my development as a researcher and genomicist. I also recognize the invaluable contributions of Dr. Sara Schaal and Dr. Laura Timm, who helped me use numerous bioinformatics programs. I am thankful for the TCU Biology Department and my fellow graduate students for working alongside me. I would also like to thank my family and friends for providing support for me during the past few years, and all the years prior as I developed an understanding of my career interests.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Our understanding of evolutionary adaptation is expanding as new methodologies increase our power to uncover the genetic basis of phenotypic variation in response to environmental change (Bomblies & Peichel, 2022). An iconic example is the disappearance of lateral plates in three-spined stickleback (*Gasterosteus aculeatus*) after colonizing freshwater lakes (Bell et al., 2004), where subsequent studies using next-generation sequencing identified loci corresponding to lateral plate formation (Hohenlohe et al., 2010). Further comparisons across lake and stream ecotypes revealed evolution acting in parallel at the same genomic regions across multiple systems, suggesting selection acted in response to environmental differences such as predation and food resources (Rennison et al., 2019). Adaptation is generally thought to span hundreds or thousands of generations. However, recent evidence shows that it can also occur on much shorter timescales (Rudman et al., 2022; Van't Hof et al., 2011). Recently, large scale anthropogenic impacts are driving environmental changes at an unprecedented rate, which will continue to result in population declines, especially species with low phenotypic plasticity or the inability to quickly adapt (Bonamour et al., 2019; Henson et al., 2017; Hoffmann & Sgrò, 2011; Malhi et al., 2020). Solutions for declining populations often involve increased human intervention to boost the number of individuals, yet such interventions may introduce their own risks (Snyder et al., 1996).

Domestication is a human-mediated intervention that can cause relaxation of natural selection and introduce artificial selection pressures to the captive environment (Balon, 2004; Christie et al., 2012; Mignon-Grasteau et al., 2005). Domestication often involves deliberate artificial selection, such as increased milk production in dairy cattle (Flori et al., 2009), coat color in domestic pigs (Fang et al., 2009), and resistance to diseases in aquaculture facilities (Hillestad et al., 2020). On the other hand, inadvertent selection can occur when the captive

environment causes genetic shifts, such as dogs ability to digest starch due to diet shifts (Arendt et al., 2016). Both deliberate and inadvertent domestication can result in divergence from their wild counterparts, and an increase in frequency of artificially selected traits may also result in unintended fitness reductions (Doublet et al., 2019; Tillotson et al., 2019).

Pacific salmon populations are generally declining across their range (Gustafson et al., 2007), and these declines are partly attributed to habitat alteration and climate change (Ainsworth et al., 2011; Crozier et al., 2021; Jones et al., 2020; Muñoz et al., 2015; Neuswanger et al., 2015; Thompson et al., 2019). Although Pacific salmon have been shown to adapt to changing environments (Fraik et al., 2021; Gilbert, 2012), numerous populations are endangered and at risk of extirpation (Gustafson et al., 2007). Hatcheries have been used to supplement declining wild populations and enhance stocks for harvest (Amoroso et al., 2017). Different from most methods of captive breeding, captive-reared salmon are released from hatcheries once they reach a certain development stage. Some hatcheries have seen increased redd abundance and spatial distribution after hatchery supplementation (Fast et al., 2015). However, rearing conditions in the hatchery may also promote divergence between hatchery and wild fish. For example, hatcheries offer a relaxed selection regime with little predation and abundant food. As a result, hatchery fish show increased competitive behavior and dominance (Metcalfe et al., 2003), changes in run timing (Ford et al., 2006), faster growth (Blouin et al., 2021; Fleming & Einum, 1997), and reduced predator avoidance behaviors compared to wild fish (Álvarez & Nicieza, 2003). Additionally, when hatchery fish are released into the wild, they generally have reduced reproductive success (O'Sullivan et al., 2020; Thériault et al., 2011) and decreased survival rates (Beamish et al., 2012; Blouin et al., 2021; Christie et al., 2012). Such divergence poses a risk to wild populations if maladapted, captive-origin individuals interbreed with wild individuals (Besnier et al., 2022; Bradbury et al., 2022; Fleming & Einum, 1997; Grant, 2011; Hagen et al.,

2019; Thériault et al., 2011; Utter, 1998). Therefore, increased knowledge on the genetic impacts of domestication could greatly benefit hatchery and wild populations.

Past studies have aimed to identify the genetic basis of domestication in response to aquaculture and hatchery rearing. Domesticated salmonids tend to be selected for early maturation, increased growth rate, and higher tolerance of parasites (Ayllon et al., 2015; Barson et al., 2015; Larsen et al., 2019). Genetic studies of farmed Atlantic salmon found differentially expressed genes, such as those associated with lipid metabolism in response to diet (Jin et al., 2020), and some expression tendencies were found in parallel across aquaculture facilities (Roberge et al., 2006). Additionally, divergent SNPs were discovered in populations of farmed and wild Atlantic salmon, yet there is little overlap in identified loci across studies (Karlsson & Moen, 2010; Karlsson et al., 2011; López et al., 2019; Mäkinen et al., 2015; Naval-Sanchez et al., 2020; Vasemägi et al., 2016; Vasemägi et al., 2012; Yáñez et al., 2016). Research on the epigenome (Gavery et al., 2018; Le Luyer et al., 2017; Leitwein et al., 2021), transcriptome (Christie et al., 2016), and genomes of Pacific salmon (Waters et al., 2018; Waters et al., 2020) also found differentiation between hatchery and wild populations. Waters et al. (2015 & 2018) used RAD-seq to identify divergent loci in Chinook salmon over four generations of hatchery rearing. Notably, the hatchery line that was segregated from the wild population showed greater genetic divergence from the wild population than the hatchery line that was integrated with wild fish, suggesting that hatchery management practices can impact the degree of genetic divergence from the wild populations (Waters et al., 2015; Waters et al., 2018). Recently, a whole-genome sequencing study on Chinook salmon showed divergence between hatchery and natural origin fish over one generation (Ford et al., 2022). However, we do not know if the genetic pathways of domestication and their fitness consequences are consistent across hatcheries. While few studies

have investigated domestication selection on a genomic level, none have done so in parallel across multiple hatchery-wild population pairs.

In this study, we further explored domestication selection in Chinook salmon using low coverage whole genome sequencing (lcWGS) in three hatchery lines of Chinook salmon. Each line was compared to its wild progenitor population within Southeast Alaska (SEAK). The three hatchery lines were separated for five to seven generations (approximately 30 years) from their wild progenitor populations. These pairwise comparisons were used to (i) discover genomic regions of differentiation across each hatchery population compared to its wild progenitor population and (ii) identify if there were parallel, shared regions of adaptive divergence across the three hatchery-wild pairwise comparisons. These results provide fine-scale genomic evidence for domestication and highlight the need to assess if certain management practices, such as integration of wild broodstock, can universally mitigate genetic risks despite multiple pathways of domestication.

Methods

Population descriptions

The National Oceanic and Atmospheric Administration's Little Port Walter Research Station (LPW) is located on southeastern Baranof Island, approximately 140 km south of Juneau (Figure 1). Since 1976, LPW has maintained a salmon research hatchery comprising two stocks of spring Chinook salmon. The focal stock for this study is derived from the Unuk River, located near Ketchikan, Alaska, which had an average annual escapement of approximately 1,800 adults in the past ten years (Meredith et al., 2022). Wild broodstock from the Unuk River was collected annually from 1976-1981 to initiate the LPW research hatchery stock (128 females and 119 males total; Templin, 2001). Wild gametes from nine males and nine females were also infused

into the LPW research hatchery stock in 1998. Each year, LPW released an average of 107,000 tagged Unuk smolts and collected all returning adults to propagate the next generation and address a variety of ecological and evolutionary questions. Only tagged LPW fish from the Unuk stock were used to spawn the following generation (i.e. segregated hatchery program), and the matings usually entailed evenly splitting the eggs from one female and fertilizing by two males.

Whitman Lake Hatchery is a production-focused hatchery located in Ketchikan, Alaska (Figure 1) and operated by the Southern Southeast Regional Aquaculture Association, Inc. The stock of spring Chinook salmon produced by Whitman Lake is derived from the Chickamin River, also near Ketchikan, which had an average annual escapement of approximately 2,000 adults in the last ten years (Meredith et al., 2022). The hatchery stock was initiated in 1981 when hatchery-origin Chickamin eggs were transferred from another facility in the region (Templin, 2001). Whitman Lake received additional hatchery-origin Chickamin eggs and fry from other facilities in 1987, 1993, 1994, and 2013. Wild broodstock was also collected annually from 1983-1987 (204 females and 104 males total; Templin, 2001). An average of 1.2 million Chickamin stock smolts have been released annually from Whitman Lake and other remote sites over the past 10 years, with approximately 13% of the fish coded-wire tagged (RMIS). The facility collects gametes from returning adults to produce the next generation, however the origin for a majority of the adults cannot be determined since only a fraction of the released fish are tagged. Therefore, there is the possibility that stray wild and hatchery fish from other stocks are occasionally spawned.

Macaulay Hatchery is a production-focused hatchery operated by Douglas Island Pink and Chum, Inc. (DIPAC) in Juneau, Alaska (Figure 1). Andrew Creek spring Chinook salmon is the wild progenitor population of the Macaulay hatchery. A tributary of the lower Stikine River, Andrew Creek had an average annual escapement of 690 adults over the past 10 years (Salomone

et al., 2022). The Andrew Creek hatchery stock was initiated at another facility from 1976-1980 when gametes were collected annually from wild broodstock (Templin, 2001). Hatchery-origin eggs and juveniles were then transferred from 1987-1992 to initiate production at Macaulay (formerly known as Gastineau Hatchery). The facility collects gametes from returning adults to produce the next generation, although additional inputs of Andrew Creek hatchery stock from other facilities have been received by Macaulay in some years. An average of 834,000 Andrew Creek stock smolts have been released annually from Macaulay and nearby locations over the past 10 years, with approximately 14% of the fish coded-wire tagged (RMIS). Similar to Whitman Lake hatchery, the origin for a majority of the adults returning to Macaulay cannot be determined since only a fraction of the released fish are tagged. Therefore, there is the possibility that stray wild and hatchery fish from other stocks are occasionally spawned. Matings at Macaulay are typically one female fertilized by two to four males.

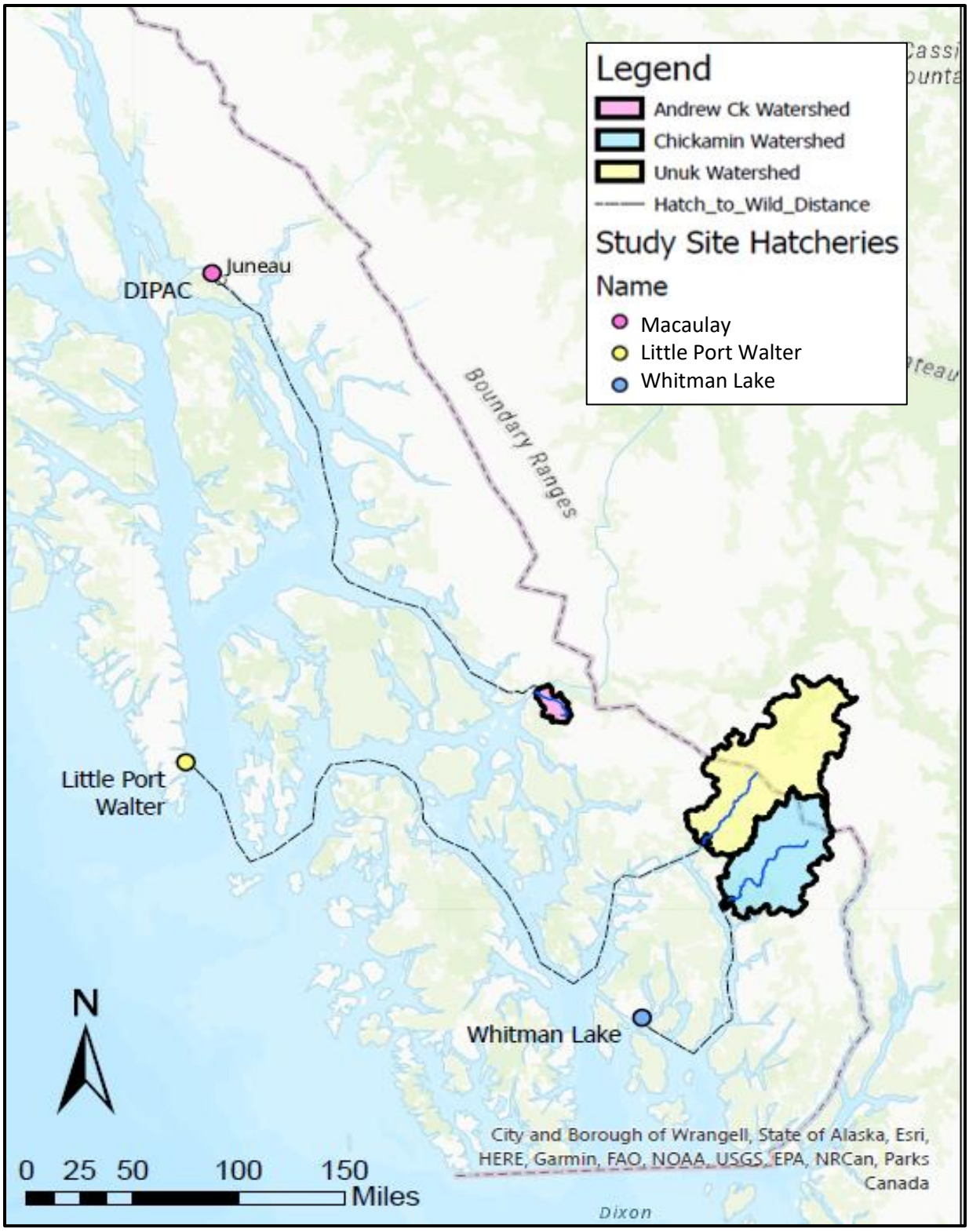


Figure 1: Site map of Southeast Alaska and the corresponding locations for each hatchery population (point) and wild population (shown as watershed). Colors represent the hatchery-wild comparison populations.

Sample collection

Fin clips were collected from returning adult Chinook salmon at hatchery facilities, and samples from wild populations were collected during spawning ground surveys by staff from the Alaska Department of Fish and Game (ADF&G). For this study, wild Unuk River samples (Unuk-W) were used from collections in 1988 and 2004, and Little Port Walter samples (Unuk-H) were collected in 2020. Phenotypic data was collected for all Unuk-H fish returning to LPW (the only population in this study with individual phenotypic data) including weight, length, sex, and age of return. Most individuals returned at age four (56%), followed by those of age five (38%), while few were age three (3%) or age six (3%). From these proportions, age of return was assumed to be four or five years; therefore, the number of generations of hatchery rearing since they were derived from the progenitor stock ranges from at least four to no more than 11 generations. The wide range is due to the infusion of wild gametes from nine males and nine females in 1998. Wild samples from the Chickamin River (Chickamin-W) were collected in 1990 and 2005, and the corresponding hatchery samples at Whitman Lake (Chickamin-H) were collected in 2014. Since the average age of return is unknown, the number of generations of hatchery rearing for Chickamin stock at Whitman Lake hatchery ranges from five to nine generations assuming a return age of four to six. Wild samples from Andrew Creek (Andrew-W) were collected in 1989 and 2004, and the corresponding samples at Macaulay hatchery (Andrew-H) were collected in 2014. Since 58% of Andrew Creek Chinook return at age five (ages four and six each represent approximately 20% of returns), a generation time of five years results in approximately eight generations of hatchery rearing since derivation from the progenitor stock (Lorna Wilson (ADF&G), pers. communication). DNA from tissue samples was extracted with Qiagen DNAeasy Blood and Tissue Extraction kits using manufacturer's protocols (Hilden, Germany).

Effective population size and population metrics from GT-sequencing

To assess the potential for allele frequency changes due to genetic drift, effective population size of the six populations was estimated using a 299 SNP panel of Genotyping-in-Thousands by sequencing (GT-sequencing) markers. The 299 SNP panel is commonly used for population assignment of Chinook from SEAK. The panel was pared down to 254 following filtering for poorly genotyped samples and SNPs with high linkage disequilibrium (LD). Using NeEstimator v2.1 (Do et al., 2014), effective population size (N_e) was determined using the linkage disequilibrium method, with a critical value set to 0.05 to remove rare alleles. Furthermore, GenAlEx (Peakall & Smouse, 2012) was used to determine observed heterozygosity (H_O). Weir and Cockerham's F_{ST} estimates were calculated using GenePop (Raymond & Rousset, 1995), and F_{ST} values were further used to conduct a principal coordinate analysis across the six populations using standardized distance-based covariance. These metrics were subsequently compared to the whole-genome sequencing metrics (i.e., F_{ST} and H_O).

Whole genome sequencing

Library prep for whole genome libraries was conducted following methods in Euclide et. al. (in prep). Samples were sent to Novogene (Sacramento, CA) for whole genome sequencing (WGS) using paired-end 150-bp reads on an Illumina NovaSeq with an intended genome-wide depth of coverage of 3x.

Sequence alignment and genotype likelihood estimation

Fastq reads were aligned to the Chinook salmon reference genome (Otsh_v1.0; GFA_002872995.1; Christensen et al., 2018) using *bwa mem* with default parameters (Li & Durbin, 2009). The aligned reads were processed with SAMtools and converted to sorted bam

files using default parameters. Then, ANGSD v0.930 (Korneliussen et al., 2014) was used to call SNPs, and genotype likelihoods were determined with the SAMtools model (GL 1) for all 192 individuals. For each SNP call, the minimum minor allele frequency was set at 5% (minMaf 0.05), and a *p*-value cutoff of 10^{-10} was used to remove rare alleles and low confidence SNPs (snp_pval 1e-10). The minimum number of individuals with genotype likelihoods at a polymorphic locus was set to 70% of the total (minInd 134), minimum depth of coverage was set to the total number of individuals (setminDepth 192), and maximum depth was set to the total number of individuals multiplied by twice the coverage, which was set to four to account for individuals with greater coverage (setmaxDepth 1500). Genotype likelihoods with at least a 99% base call accuracy (minQ 20) and mapping accuracy (minMapQ 20) were retained. Major and minor alleles for all individuals were determined from genotype likelihoods (doMajorMinor 1).

To explore genetic divergence across populations, principal component analyses (PCAs) were conducted using PLINK v1.9 (Purcell et al., 2007). The corresponding covariance matrix led to the identification and removal of four individuals (three from Unuk-H, one from Chickamin-W) that skewed the clustering, which may have been due to relatedness and missing data. All further analyses utilized the remaining 188 individuals. Further PCAs were conducted using genotype likelihoods with PCAngsd (Meisner & Albrechtsen, 2018). To ensure there was no population structure between wild samples efforts, individuals were also identified by sampling effort. Due to lack of clustering between efforts, wild samples were combined into one population.

Identification of regions with high genomic divergence

To determine weighted pairwise F_{ST} (Weir and Cockerham's) for the three hatchery-wild pairs, site allele frequency likelihoods were calculated in ANGSD (doSaf 1) using the same filter

criteria as above for each population, except the SNP *p*-value cutoff was set to 10^{-6}. Using the folded site frequency spectrum for each hatchery-wild pair (realSFS), global (realSFS fst stats) and genome-wide F_{ST} (realSFS fst stats2) was calculated. Manhattan plots of genome-wide F_{ST} were plotted in R on a per-SNP basis to visualize genetic differentiation across hatchery and wild pairs.

We identified F_{ST} peaks of interest to investigate genomic regions that may be responding to domestication selection within each hatchery-wild population pair. First, only the top 5% of SNPs with the highest F_{ST} values across the genome were retained for sliding window analyses to remove background noise of non-outlier F_{ST} SNPs. The genome was then divided into non-overlapping windows of sizes 10 kilobases (KB) and 100 KB, which were used to capture both narrow and wide peaks. For each size, windows were determined to be outlier peaks if they met two criteria: i) the maximum F_{ST} value in the window was within the top 1% of F_{ST} values from the retained SNPs, and ii) SNP density in the window was within the top 0.5% of densities across all windows. These criteria were established to find windows with numerous high-F_{ST} SNPs that ultimately created an outlier peak (these identified high F_{ST} regions are hereafter called peaks). Furthermore, peaks were compared across each hatchery-wild pair to identify differentiating regions shared across hatcheries. Chinook salmon genes were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/genome/13133?genome_assembly_id=360171), and location of genes was compared to outlier peaks, identifying those within and near windows. Further analyses on these peaks included exploration of potential structural variants using the *lostruct* R package (Li & Ralph, 2018), which created local PCAs in 50-SNP non-overlapping windows. Using the Euclidean distance output, multidimensional scaling (MDS) was plotted to discover outliers across the genome and determine if any identified peaks showed signs of a structural variant. Additionally, local PCAs were conducted within 5 KB on either side of the SNP with the

highest F_{ST} value of each outlier peak (for a total of 10 KB) to assess potential structural variation at the peaks. Furthermore, genotype likelihood heatmaps were developed using custom R scripts (originally developed by Sara Schaal) across the same 10 KB regions as the local PCAs to visualize allele frequency differences within peaks.

Linkage disequilibrium (r^2) was calculated in ngsLD (Fox et al., 2019) within each identified outlier peak. Within each outlier peak, LD was calculated for all SNP pairs within the same 10 kb regions as described above for genotype likelihood heatmaps. LD was calculated within each population separately, so that hatchery and wild LD trends could be compared to one another. Subsequently, LD was calculated across each hatchery-wild comparison by including all individuals from both hatchery and wild pairs to increase sample size for follow-up tests. To test if LD in the peaks deviated from background rates of LD, additional r^2 estimates were calculated for ten randomly selected 10 KB regions outside of outlier peaks. This was performed to help interpret whether outlier peaks were driven by selection or drift (drift would likely be responsible if LD was similar between outlier regions and the randomly selected regions, whereas selection would likely be responsible if LD was consistently higher in outlier regions). To account for potential bias due to the relationship between LD and distance, a randomization without replacement method was used to subset the background r^2 values, which was repeated 600 times. The number of background r^2 values that were subset was equal to the number of r^2 values in the corresponding outlier peak to maintain equal sample sizes. For each randomization permutation, a Wilcoxon Rank-Sum test ($\alpha = 0.05$) was performed to compare the outlier peak r^2 values to the background r^2 values. If the same statistical pattern (i.e., significant difference in r^2 between peaks and background) was found in at least 90% of the permutations then the difference in r^2 values was determined to be significant.

Additional genome-wide analyses were conducted to determine if F_{ST} outlier peaks were supported by other metrics. Absolute genetic divergence (D_{xy}) was calculated on a per-SNP basis using ngsTools' *getDxy.pl* script (Fumagalli et al., 2014) to further explore the modes of selection occurring at these peaks. Subsequent analyses of Tajima's D were calculated in ANGSD (thetaStat do_stat) with a sliding window of 10,000 bases and a step of 5,000 bases. Similarly, genome-wide heterozygosity was calculated on a per-SNP basis using ngsTools' ngsStat (Fumagalli et al., 2013) to investigate genetic diversity within populations and compare heterozygosity across hatchery and wild population at each peak. Using custom scripts in R, heterozygosity was plotted using non-overlapping windows consisting of 15,000 bases. Heterozygosity was further explored by determining if it was significantly different between the wild and hatchery populations within each outlier peak using a Wilcoxon Rank-Sum test ($\alpha = 0.05$).

Results

Low coverage genomic sequencing produced an average of 63 million reads across each of the 188 individuals in the WGS dataset. The average percent coverage for each base pair was 78%, and the average depth of coverage was 3.6X (range = 1.4X – 6.8X). After quality filtering and SNP scoring, the final set of retained genotype likelihoods for each population averaged 6 million SNPs (range = 4,658,433 – 7,034,208).

Principal component analyses across all individuals, comprising 1.1 million SNPs, allowed for visualization of the population structure. The first and second principal components explained 1.62% and 1.30% of the variance, respectively (Figure 2A). Samples from Chickamin-H and Chickamin-W overlapped the most out of all hatchery-wild pairs, followed by Andrew-H and Andrew-W. However, the third principal component accounted for variation (0.62%)

between Chickamin-H and Chickamin-W (Figures A1A & A1B). Unuk-H was the most dispersed and did not overlap with Unuk-W, suggesting a high degree of genetic difference between Unuk-H and the Unuk-W relative to the other population comparisons. Unuk-W and Chickamin-W were adjacent in the PCA, which is reflective of their geographic proximity (see Figure 1). One individual from Chickamin-W and another from Andrew-W clustered near Unuk-H, which was unexplained and may suggest straying from other hatchery populations into Andrew-W and Chickamin-W (both had depth of coverage greater than four, suggesting this was not due to poor sequencing). The PCA results were further corroborated from the principal coordinate analysis (PCoA) using GT-seq data at 254 SNPs for each population, which showed similar clustering patterns (Figure 2B).

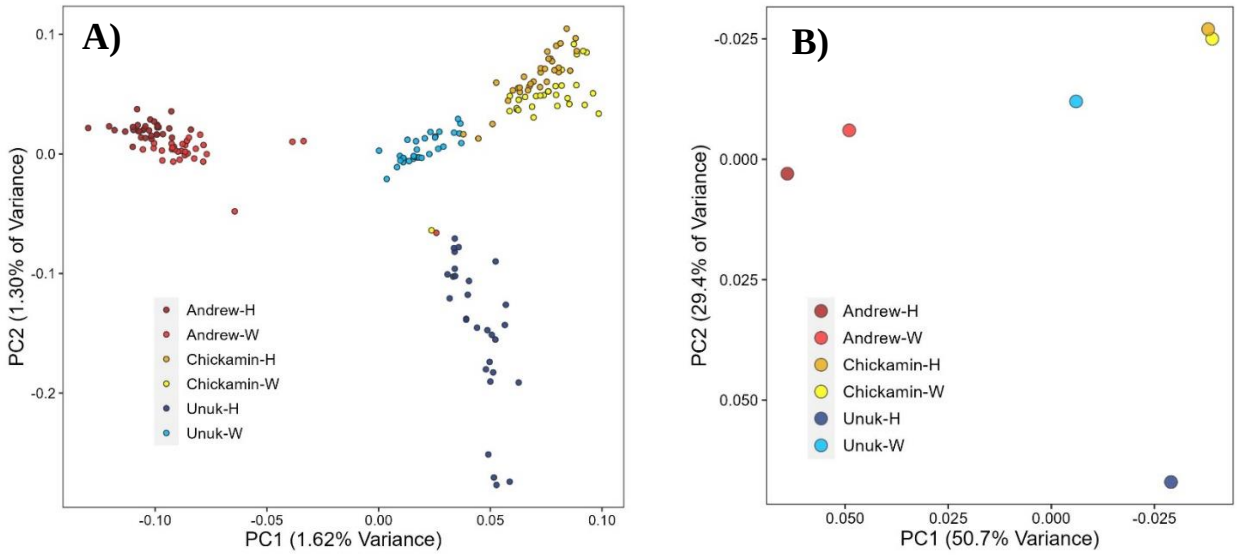


Figure 2: Principal component analysis (PCA) from (A) whole-genome sequencing data for each sample, and (B) principal coordinate analysis (PCoA) from GT-sequencing data for each population.

Estimates of effective population size (N_e) served as proxies of genetic diversity within each population, where Unuk-W was 15 times that of Unuk-H, which had the lowest N_e of all populations (1,397 vs. 92, respectively, Table 1). Andrew-W had the largest N_e (1,582), which

was approximately three times greater than Andrew-H (535). In contrast to the other two wild populations, Chickamin-W was 1.3 times greater than Chickamin-H and was also less than the N_e at Andrew-H. Observed heterozygosity from WGS and GT-seq did not vary greatly across populations, although H_O was slightly less for GT-seq calculations compared to WGS (Table 1). Interestingly, whenever H_O from WGS was greater in the wild population compared to the corresponding hatchery population, H_O from GT-seq was greater in the hatchery population, although the differences between the two are minor (maximum difference between hatchery and wild H_O per population pair is 0.009).

Table 1: Hatchery and wild population pairs, sample sizes for WGS and GT-sequencing samples with their collection year, estimated effective population sizes (N_e) with corresponding parametric confidence intervals, and observed heterozygosity (Ho).

Site	**WGS n (Year)**	**GT-seq n (Year)**	**GT-seq N_e (CI)**	**WGS Ho (SE)**	**GT-seq Ho (SE)**
Unuk-H	32 (2020)	46 (2018)	92 (78 - 109)	0.295 (0.00006)	0.280 (0.012)
Unuk-W	16 (2004) 16 (1988)	91 (1988)	1397 (679 - Infinity)	0.298 (0.00006)	0.277 (0.012)
Andrew-H	32 (2014)	46 (2014)	535 (288 - 2895)	0.299 (0.00005)	0.269 (.012)
Andrew-W	16 (2004) 16 (1989)	188 (2004)	1582 (1010 - 3485)	0.296 (0.00005)	0.272 (.011)
Chickamin-H	32 (2014)	47 (2014)	264 (188 - 441)	0.302 (0.00006)	0.270 (0.012)
Chickamin-W	16 (2005) 16 (1990)	91 (2005)	349 (274 - 477)	0.296 (0.00006)	0.279 (0.012)

Global F_{ST} estimates between all populations revealed low genetic differentiation between the populations. The GT-seq estimates were generally similar to WGS estimates, and F_{ST} estimates were greatest between Andrew-H and Unuk-H for both methods (GT-seq F_{ST} = 0.0291; WGS F_{ST} = 0.0231). In some cases, WGS data had greater values than GT-seq, such as in the Chickamin H/W comparison (GT-seq F_{ST} = 0.0079; WGS F_{ST} = 0.0102). Out of the three

hatchery-wild pairs, Andrew had the lowest combined estimates (GT-seq F_{ST} = 0.0033; WGS F_{ST} = 0.0088), whereas Unuk had the greatest differentiation (GT-seq F_{ST} = 0.0184; WGS F_{ST} = 0.0138), which aligns with visualized genetic distances from the PCA and PCoA.

Table 2: Pairwise global F_{ST} across all loci with WGS (below the diagonal) and GT-seq (above the diagonal). Darker red colors represent greater F_{ST} values. Bold values are hatchery-wild population pair comparisons.

	Unuk-H	**Unuk-W**	**Andrew-H**	**Andrew-W**	**Chickamin-H**	**Chickamin-W**
Unuk-H	0	**0.0184**	0.0291	0.0263	0.0220	0.0211
Unuk-W	**0.0138**	0	0.0162	0.0117	0.0118	0.0080
Andrew-H	0.0231	0.0162	0	**0.0033**	0.0284	0.0268
Andrew-W	0.0201	0.0137	**0.0088**	0	0.0202	0.0217
Chickamin-H	0.0183	0.0113	0.0222	0.0204	0	**0.0079**
Chickamin-W	0.0176	0.0111	0.0220	0.0195	**0.0102**	0

The genome-wide Manhattan plots further identified genomic regions with greater F_{ST} than the background across each hatchery-wild pair (Figure 3). The Chickamin H/W comparison showed the least genome-wide differentiation, whereas global F_{ST} suggested Unuk-H and Unuk had the greatest levels of differentiation. In all populations, some regions contained numerous SNPs that formed identifiable peaks, which were approximately five to ten kilobases wide. Two peaks in the Unuk hatchery-wild comparison on chromosomes four and nine were particularly pronounced (Chr. 4: Average F_{ST} = 0.173, Max F_{ST} = 0.370; Chr. 9: Average F_{ST} = 0.148, Max F_{ST} = 0.337). The quantifiable method used to determine outlier peaks led to the discovery of 17 peaks in Unuk, 15 peaks in Andrew, and five peaks in the Chickamin hatchery-wild comparison (see Table A1). None of the peaks were located in the same genomic region across the hatchery-wild pairs.

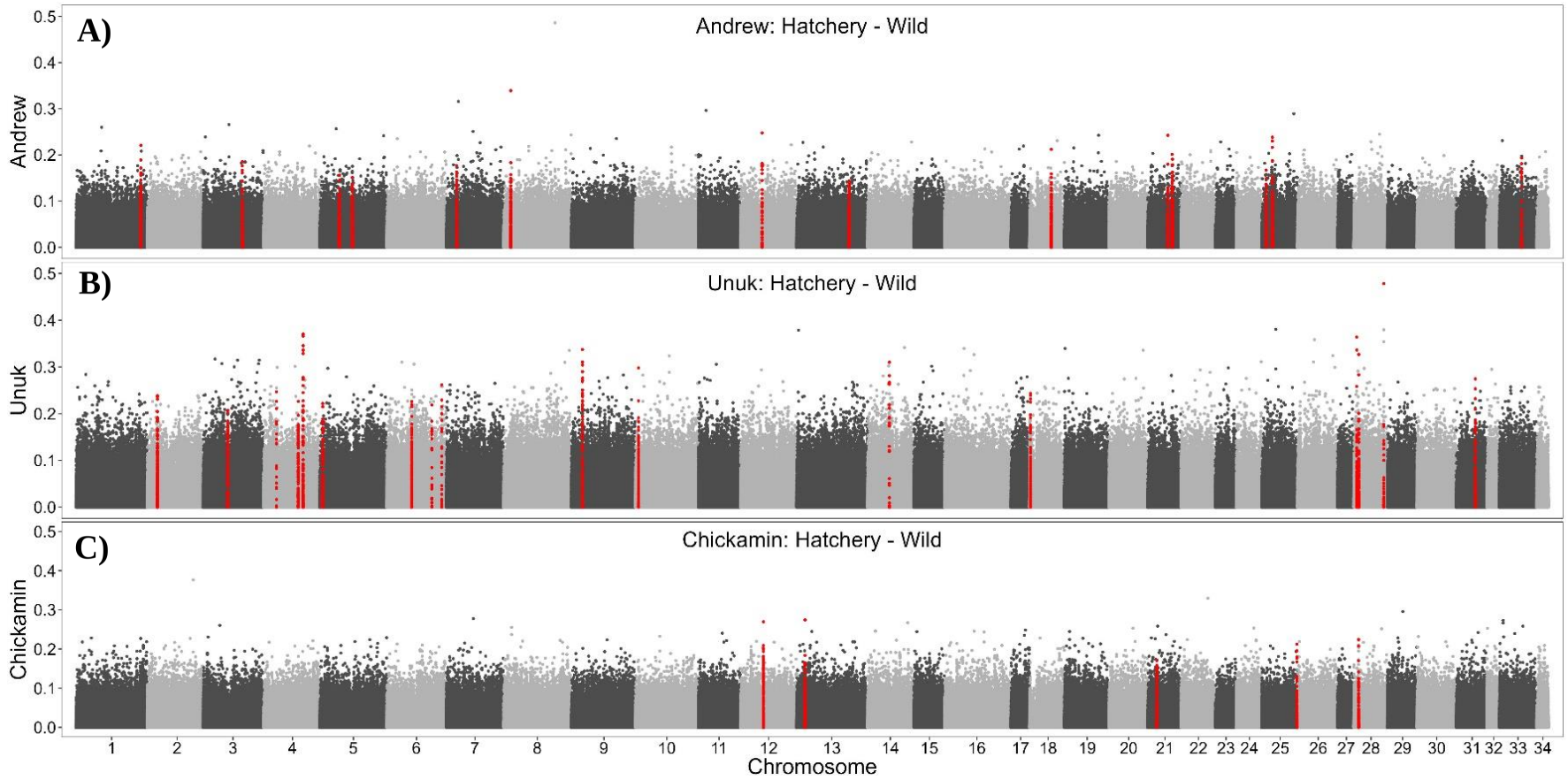


Figure 3: Genome-wide Manhattan plots displaying F_{ST} between hatchery-wild population pairs; (A) Andrew-H/Andrew-W, (B) Unuk-H/Unuk-W, (C) and Chickamin-H/Chickamin-W. Alternating grays represent different chromosomes, and red SNPs are those within identified outlier peaks.

Peaks were better visualized by zooming in on the regions of interest (Figure 4a). Numerous genes overlapped with or were close to the peaks, but no gene ontology analysis was conducted, so the function is still unknown (Table 3). Additionally, MDS outliers from *lostruct* analyses did not fall in the same regions as outlier F_{ST} peaks (Figure A4). After conducting PCAs in the MDS outlier regions, there was no distinct clustering into two or three groups, as would be expected for structural variants (Huang et al., 2020). Furthermore, local PCAs for SNPs within each of the outlier peaks also showed little to no sign of separation into two or three clusters, suggesting that structural variants across hatchery-wild populations is not likely.

Absolute genetic divergence (D_{xy}) was generally higher in F_{ST} peaks, such as in the Unuk outlier peak on chromosome four (Figure 4C), although there were numerous peaks that had slight or no elevation in D_{xy}. The genotype likelihood heatmap for chromosome four corroborated D_{xy} values since the minor allele became the major allele in the hatchery population, which shows a departure from the ancestral allele (Figure A3). This was delineated in numerous genotype likelihood heatmaps for other peaks, but not all showed an increase in D_{xy} (Table 3). For the peak at 53 MB on chromosome four, heterozygosity was significantly greater in the Unuk-W population compared to Unuk-H ($p = 0.0125$), although heterozygosity averaged across windows, as shown in Figure 4C, does not seem dramatically different. Tajima's D is slightly elevated within the same, but it also does not seem dramatically different from background deviations (Figure 4E).

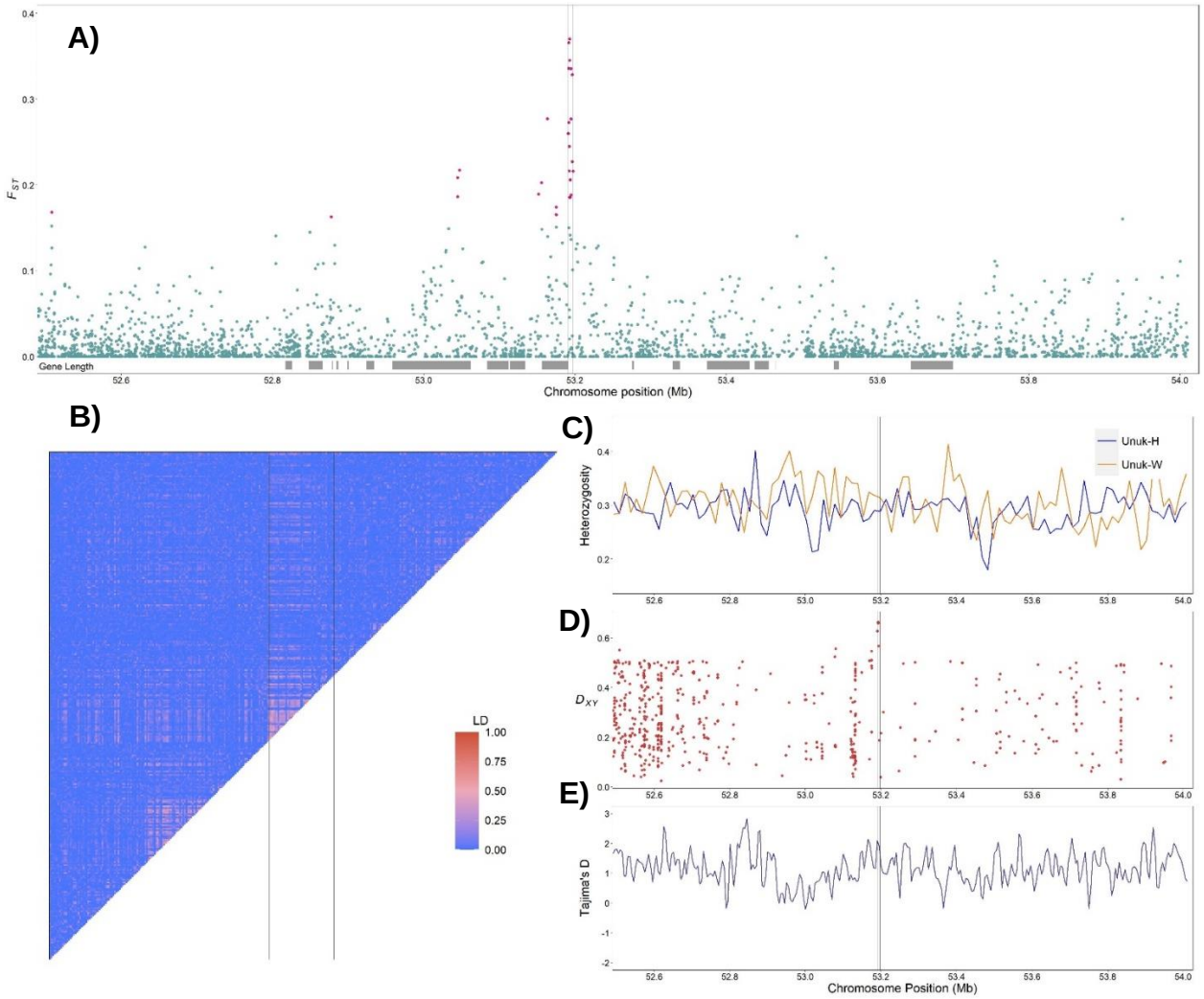


Figure 4: Unuk H/W on chromosome 4 outlier peak around 53 Mb. (A) Manhattan plot of F_{ST} outlier peak, showing genes that are within or near the peak, (B) LD (r^2) heatmap including SNPs for hatchery and wild populations within the same peak across a total of 10 KB – red is higher LD and blue is lower LD; (C) Heterozygosity – yellow is hatchery, blue is wild; (D) D_{xy} for Unuk H/W; (E) Tajima's D for the Unuk-H (LPW) population. Gray vertical lines represent the outlier peak window, which is 6.6 KB wide.

Linkage disequilibrium heatmaps revealed regions of high LD within the peaks for the combined hatchery and wild populations (Figure 4B). To parse out the LD contributions of the hatchery and wild populations, they were separately plotted as LD heatmaps within each peak. These results showed that the patterns of LD were not consistently shared between peaks (as shown in Figure A3). In some peaks, LD was greater in the hatchery population; in others, it was greater in the wild population. For example, the peak on chromosome 9 of Unuk H/W had

greater LD in the hatchery population than in the wild population but note that no statistics were calculated (Figures S2C & S2D). Overall, most peaks showed similar levels of LD across both the hatchery and the wild populations. However, LD within combined hatchery and wild populations was generally elevated in outlier peaks compared to background levels of LD (Figure 5). Thirty-three of the 37 outlier peaks had significantly greater LD than background levels (Table 3). However, the most visually identifiable F_{ST} peak from the Chickamin comparison did not show a greater LD than the background (Figure 5C), which was surprising since the LD heatmap had high r^2 values in the peak (Figure A5). This may suggest that LD is generally greater on chromosome 12 than other chromosomes, or it may be a function of random sampling.

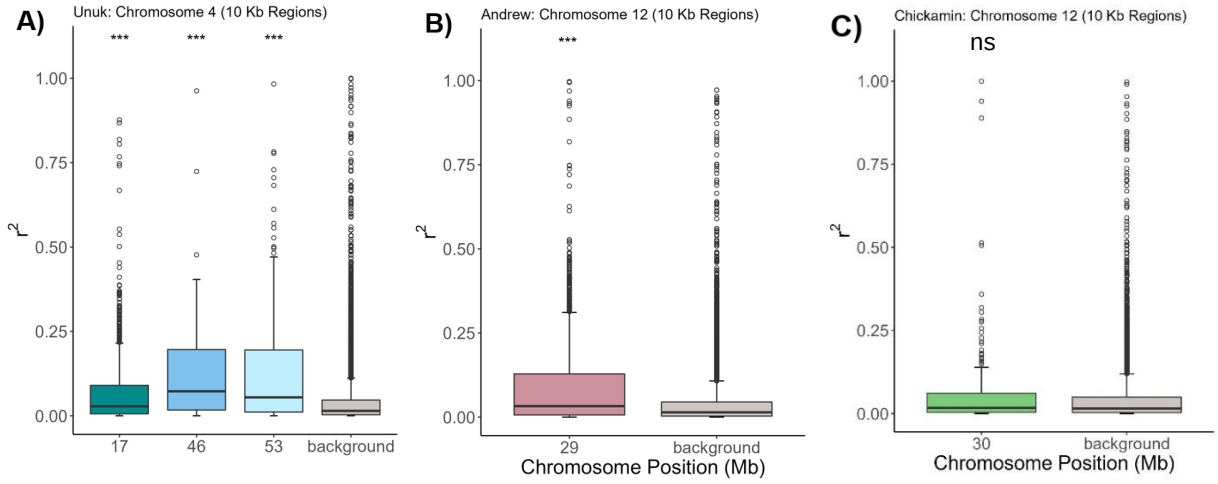


Figure 5: Boxplots of LD (r^2) calculated for hatchery & wild samples of five of the 37 outlier peaks on (A) Unuk chromosome 4, (B) Andrew chromosome 12, and (C) Chickamin chromosome 12. Outlier peaks were statistically compared to background LD on the same chromosome of each peak. *** p-value < 0.0001, ns = not significant.

Additional metrics were determined across all outlier peaks. Overall, most peaks did not have differing heterozygosity between hatchery and wild populations (Table 3). Across all comparisons, heterozygosity was greater in the hatchery population at six peaks and also greater

in the wild population at another six peaks, whereas the remaining 25 peaks had no statistical difference across hatchery and wild populations (Table 3).

Table 3: Number of outlier peaks per comparison and additional descriptions of the peaks including maximum F_{ST}, average size in kilobases, number of genes within peaks, and the total peaks that had significantly greater LD (r^2). Elevated LD in peaks was statistically measured compared to background LD using a Wilcoxon Rank Sum Test (Bonferroni correction; $\alpha = 0.05$). Statistically different observed heterozygosity (Ho) in either the wild or hatchery population within outlier peaks was determined using a Wilcoxon Rank Sum test ($\alpha = 0.05$), where some peaks had greater heterozygosity in the hatchery population and some greater in the wild (see Table A1 for peak-specific metrics).

Population Pair	Number of Peaks	Max F_{ST} in Peaks	Average Size of Peaks (Kb)	Number of Genes	Signif. Elevated LD	Signif. Different Ho
Unuk-H Unuk-W	17	0.478	34.5	25	14	3 5
Andrew-H Andrew-W	15	0.339	33.3	17	12	2 1
Chickamin-H Chickamin-W	5	0.274	20.6	8	4	1 0

Discussion

Many studies have observed a rapid reduction of fitness in multiple hatchery lines of Pacific salmon (Blouin et al., 2021; Christie et al., 2014; O'Sullivan et al., 2020). There have been numerous mechanisms proposed that drive these differences, but the underlying effects and relative importance of domestication selection has not been fully elucidated (Gavery et al., 2018; Le Luyer et al., 2017; Mäkinen et al., 2015). In this study, we investigated Chinook salmon in three independent hatchery lines and compared them to their wild progenitors. We used whole genome sequencing to discover outlier regions of differentiation and compared these regions to determine if they were conserved across hatcheries. We found signatures of domestication selection following approximately six to eight generations of hatchery rearing. However, the degree of differentiation varied depending on the hatchery line. When utilizing outlier peaks to distinguish regions of potential domestication selection, no regions were found in parallel across

all three hatchery-wild population pairs, suggesting the targets of domestication selection vary between the three hatchery lines.

Population divergence and diversity

The Chickamin and Andrew hatchery-wild F_{ST} estimates were the least differentiated of all F_{ST} comparisons. This, along with the principal coordinate analysis results (see Figure 2), suggests that these two hatchery lines have less differentiation from their wild progenitor populations relative to Unuk, although note that PC3 shows genetic distance between Chickamin-H and Chickamin-W (Figure A1). Theoretically, when comparing wild populations to one another, the differences are attributed to standing genetic variation across stocks; when comparing hatchery populations to one another, the differences are attributed to both standing genetic variation from their founding population and the effects of domestication selection. Therefore, it is expected that the wild progenitor population would show the lowest amount of divergence to its hatchery line than any other population because it would only be accounting for changes due to domestication selection or drift. This was not the case for Unuk-H, where global F_{ST} estimates showed greater differentiation between Unuk-W and its corresponding hatchery line than to the other two wild populations. Therefore, six to eight generations of captive-breeding in Unuk-H created more differentiation from its wild progenitor population than to salmon from adjacent watersheds.

The divergence between Unuk-W and Unuk-H may largely be due to the low effective population size of 92 at Unuk-H, which was expected because the LPW Research Station has a smaller number of individuals in the experimental hatchery line than the two production-focused hatcheries (107,000 annual releases at Unuk-H compared to over 800,000 and 1.2 million at Chickamin-H and Andrew-H, respectively). Another potential reason for the low N_e in the Unuk-

H population is because LPW is likely more segregated than the other two hatchery lines due to 100% coded wire tagging and real-time broodstock screening prior to spawning, although all three hatcheries are presumed to have little to no gene flow with their progenitor population. This lack of gene flow likely contributes to a reduction in N_e, which seems to result in stronger genetic drift within Unuk-H than the other hatchery lines (Martinez et al., 2022). Although the other two hatchery populations have greater N_e than Unuk-H, there is still reduced N_e for both Chickamin-H and Andrew-H compared to their wild counterparts. Andrew-H was the only hatchery line with an N_e estimate greater than 500, which is predicted to be large enough to maintain genetic diversity in a population (Franklin & Frankham, 1998). Therefore, the effective population sizes at these hatcheries provide a broad understanding of the potential severity of genetic drift in the hatchery populations (Naish et al., 2013), with Unuk-H showing greater genetic drift than Chickamin-H and Andrew-H.

Heterozygosity differences between populations seem largely unaffected across hatchery-wild comparisons. This may allude to gene flow between the hatchery line and strays that are inadvertently incorporated in broodstock for the two production-focused hatcheries. For example, Waters et. al. (2015 & 2018) found greater differentiation from the wild population and inbreeding in the segregated line (broodstock consisting of only hatchery fish) compared to the integrated line (broodstock consisting of hatchery and wild-origin fish). Since the current study has no quantitative data on gene flow in the Chickamin-H and Andrew-H line, we cannot directly compare gene flow across hatchery lines like that of (Waters et al., 2015). However, if individuals from the progenitor stock are introduced into the hatchery lines, the signatures of selection may get lost if alleles adapted to the hatchery environment are mixed with wild alleles (Bourret et al., 2011). Similar levels of genetic diversity across domesticated and wild populations were found in Atlantic salmon (Mäkinen et al., 2015; Vasemägi et al., 2012) as well

as Pacific salmon (Smith et al., 2014; Waples et al., 1990). Moreover, heterozygosity decreases at a slower rate as N_e during a bottleneck, so the lack of reduction may be a function of few generations since the hatchery populations were founded (Allendorf, 1986; Nei et al., 1975).

Outlier F_{ST} peaks

Due to the process of hitchhiking, neutral regions that are in strong linkage with an allele under selection will also see changes in allele frequencies, resulting in a selective sweep (Hermisson & Pennings, 2005). Selective sweeps in novel environments often utilize standing genetic variation, where alleles that were unfavorable or neutral in the original environment are favored in the new environment (Hermisson & Pennings, 2005). Selective sweeps can create footprints of selection, which WGS can detect with more precision than previously used methods such as RAD sequencing (Lou et al., 2021). In this study, we quantitatively identified peaks for detection of selective sweeps across the genomes of each hatchery-wild pair. From the hatchery-wild comparisons, 17 outlier peaks were detected in Unuk, 15 in Andrew, and five in Chickamin. Hitchhiking selection is likely driving some of the allele frequency differences within many outlier peaks due to increased LD in 80% of peaks compared to background estimates of LD. Since genetic drift is expected to reduce recombination genome-wide, the increased LD in the outlier peaks potentially indicates localized selective sweeps (Ohta & Kimura, 1969; Slatkin, 2008). Generally, Tajima's D within peaks did not deviate from the background rate, although the genome-wide trend was positive. A positive Tajima's D can indicate either balancing selection or a recent decrease in population size (Tajima, 1983), the latter of which is supported from N_e estimates. Furthermore, Tajima's D may not have the power to detect the signatures of recent domestication (Innan & Kim, 2004), which is most likely the case in the present study. However, the combined metrics demonstrate that selective sweeps are likely the driving force in

most of the outlier peaks, which may be a direct effect of domestication selection (Liu et al., 2014).

Although analyses were conducted on all identified outlier peaks, we focused on the peak on chromosome four (53 Mb) in the Unuk hatchery-wild comparison. This peak is approximately 6.6 kb long, which is smaller than most peaks (average = 32.0 kb). The LD heatmap showed statistically greater LD within the peak compared to background regions of chromosome four (average p-value = $1.4x10^{-49}$). Heterozygosity within the peak was greater in the wild population than the hatchery population, averaging 0.325 and 0.285, respectively ($p = 0.015$). However, heterozygosity between wild and hatchery populations across all outlier peaks did not show consistent trends, with over two-thirds of peaks showing insignificant differences. Estimates of absolute genetic divergence **(**D_{xy}) were also elevated at this peak, which suggests the allele frequency changed such that the minor allele in the wild population increased in frequency in the hatchery population to become the major allele. Thus, for outlier peaks that corresponded to elevated D_{xy}, the hatchery population's haplotype diverged from that of the ancestral allele (supported by genotype likelihood heatmaps, Figures A3C & A3F; Han et al., 2017). Similar to Tajima's D, D_{xy} is known to have lower power than F_{ST} when the divergence time is short (Cruickshank & Hahn, 2014). As a result, the change in D_{xy} between hatchery and wild is not pronounced when comparing our estimates of D_{xy} to other studies (e.g., Clucas et al., 2019; Ravinet et al. 2018). Another reason for lower D_{xy} may be because other studies found locally adapted alleles with high gene flow, which is another mechanism that can increase D_{xy} (Han et al., 2017). One way populations can locally adapt with high gene flow is through structural variants such as inversions (Huang et al., 2020). However, the analyses presented here suggest no evidence of inversions and support selection on a locus-by-locus basis.

Evidence for domestication selection

Although the data suggest domestication selection is operating on multiple genes within the Chinook salmon genome, none of the outlier peaks were shared across hatchery-wild comparisons. Previous studies that identified parallel changes in salmon across captive-bred facilities largely identified genes with shared function (Le Luyer et al., 2017; Roberge et al., 2006). For example, in a study addressing transcriptional differences in farmed and wild Atlantic salmon, Roberge et al. (2006) found 16% of gene expression transcripts to be in parallel across aquaculture facilities. However, no specific loci were found in parallel; rather, they identified gene expression patterns that were responsible for similar functions, suggesting that the aquaculture populations took different pathways to the same result (Roberge et al., 2006). Le Luyer et al. (2017) similarly found regions of the muscle epigenome to be hypermethylated in two hatchery coho salmon (*Oncorhynchus kisutch*) populations compared to wild populations, and this hypermethylation was consistent across populations. A follow-up study found similar results of hypermethylation in the epigenome of male germ cells (Leitwein et al., 2021). Some alleles have been found in parallel across multiple aquaculture facilities, but most are not shared across populations (López et al., 2019). Overall, these studies, together with the results presented herein, suggest that domestication selection is likely polygenic and targets multiple genes involved in many different biochemical pathways.

Domestication selection may be difficult to detect in the genome during the first few generations may be difficult, especially when selection is weak and genetic heterozygosity is high (Mäkinen et al., 2015; Waters et al., 2015). This has been a common complexity of domestication selection studies in hatchery-reared or farmed salmonids, especially when studies utilized lower coverage methodologies such as RAD-sequencing and SNP chips (López et al., 2019; Mäkinen et al., 2015; Vasemägi et al., 2012). In response to the lack of evidence linking

genetic changes to phenotypic differences in captive-bred individuals, some studies have attributed fitness changes to epigenetics and transcriptomics, theorizing a genotype-by-environment response to domestication pressures in salmonids (Christie et al., 2016; Gavery et al., 2018; Le Luyer et al., 2017; Leitwein et al., 2021). While those mechanisms likely contribute to differences between hatchery-reared and wild salmon, the results of this study suggest that domestication selection is prevalent in the genomes of multiple hatchery lines. This is further supported by a recent study by Ford et al. (2023), which found genetic divergence in Chinook salmon at greater rates than would be expected at random after only one generation of hatchery-rearing. This suggests that divergence between hatchery and wild populations is prevalent, but high-resolution genomic techniques are needed to discover regions of differentiation. These regions of divergence across the genome may be undergoing allele frequency shifts in response to differences between the hatchery and wild environments. Furthermore, the identified outlier peaks may not be the only genomic regions experiencing domestication selection in these populations, and further investigation may identify quantitative traits with additive effects (da Silva Ribeiro et al., 2022).

Limitations and future directions

There may be inherent genetic differences in the three comparisons that contribute to the lack of overlap across hatchery-wild comparisons. First, the three wild populations are genetically distinct stocks, which means that there might not be the same standing genetic variation in all populations. Another possible reason for the lack of parallel signals of domestication selection could be attributed to differences in hatchery practices across sites, which may have caused different paths of domestication selection. Lastly, the wild samples were collected after the hatchery stock was established, so it is possible that either i) the wild samples

do not fully represent the genetic architecture of the founding broodstock, or ii) selection occurred in the wild populations after the hatcheries were established. However, the latter is unlikely because F_{ST} comparisons between the two wild collections years did not find outlier peaks in the same regions as those identified in the hatchery-wild comparisons (Figure A2B).

An additional limitation of this study is the lack of phenotypic data, which eliminated the option of attributing genes within peaks to phenotypic differences across wild and hatchery populations. Genome-wide association studies previously conducted in wild Chinook salmon compared to salmon from two different hatchery management regimes discovered trait-associated loci related to changes in weight, run timing, and maturation (Waters et al., 2018). However, more research is needed to further elucidate the genetic architecture of phenotypes associated with domestication selection. As whole genome sequencing becomes more accessible, future studies should address relationships between phenotypic and genetic changes in hatchery-reared salmon, which highlights the need for phenotypic data collection of hatchery broodstock.

Conclusions

Further investigation in the differences between the management of hatchery lines may elucidate why selection was seemingly stronger in the Unuk hatchery line than Andrew and Chickamin. As previously mentioned, there is potential that smaller, more segregated hatchery lines may experience stronger domestication selection than those that experience greater gene flow (Martinez et al., 2022; Waters et al., 2015). However, additional studies directly addressing different hatchery practices would be necessary to determine how operations may impact genetic diversity and fitness of hatchery salmon. Even though the regions under selection were not shared between hatchery lines, there may be similar genetic pathways under selection across locations. Regardless, these results suggest that hatchery fish are experiencing domestication

selection, and that their genomes are deviating from their wild progenitor populations after only a few generations of hatchery-rearing.

APPENDIX

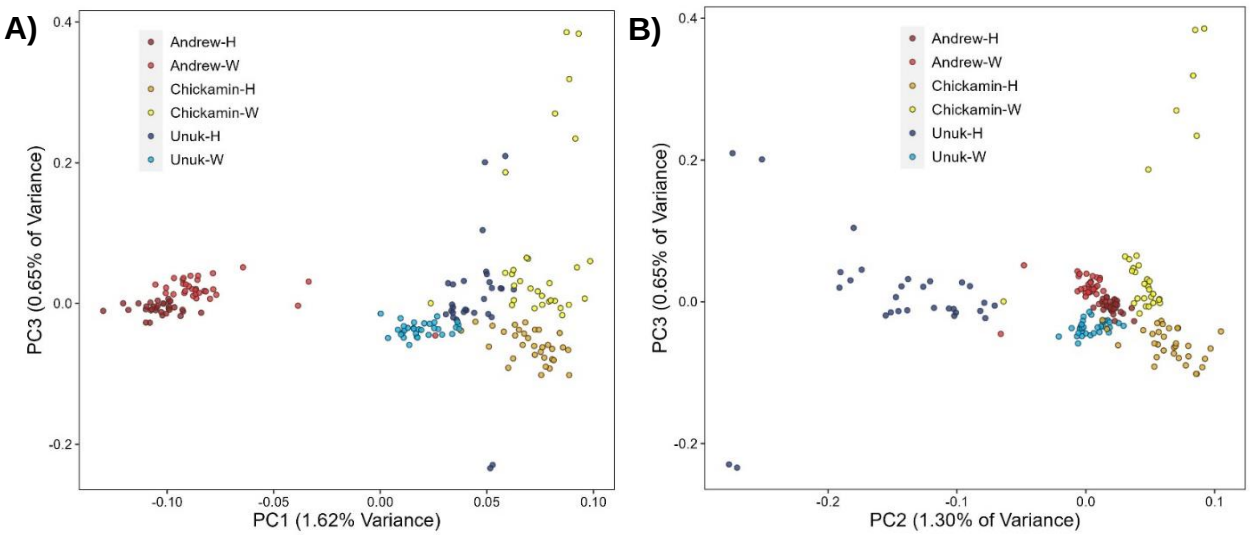


Figure A1: (A) PC1 vs. PC3 (B) PC2 vs. PC3, and for which PC3 shows spread between Chickamin-H and Chickamin-W.

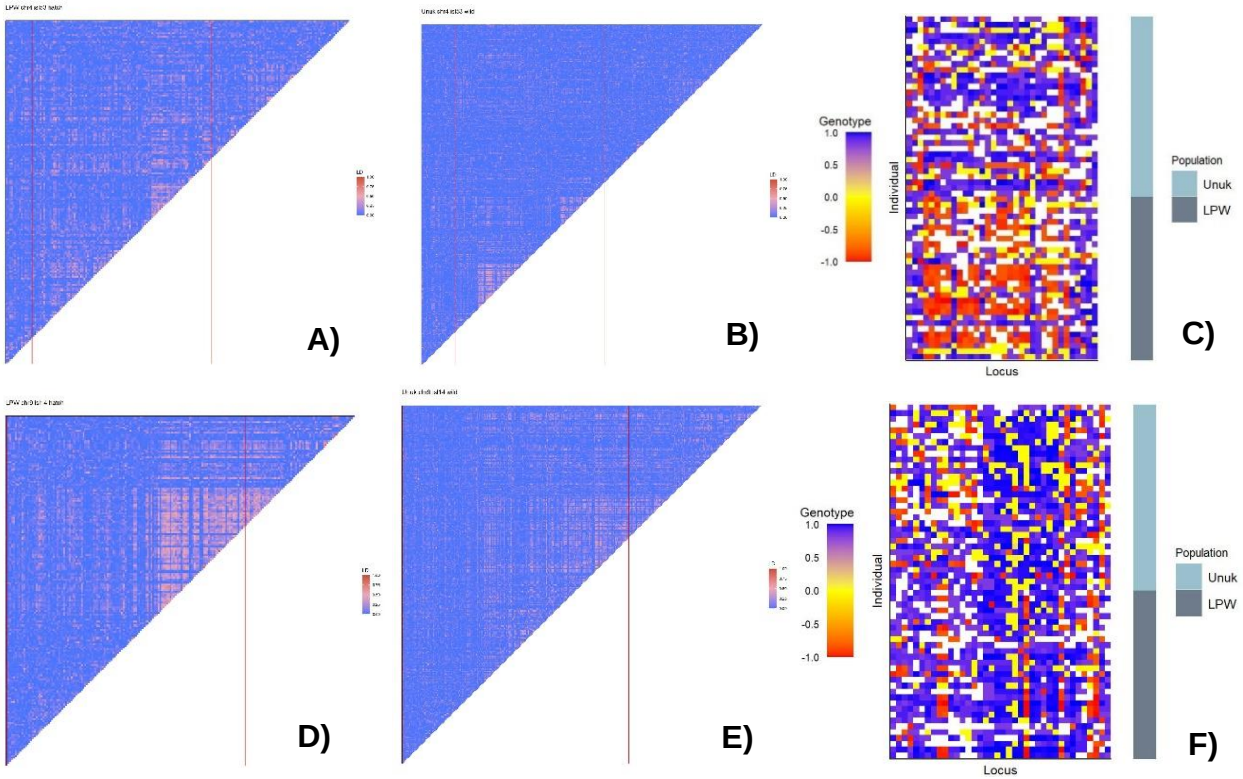


Figure A2: LD heatmaps at chromosome 4 around the peak at 53 Mb (A) only Unuk-H samples, (B) only Unuk-W samples, and (C) corresponding heatmap of genotype likelihoods for each sample (blue is reference allele, yellow is heterozygous, and red is alternative allele; the opacity represents likelihood of correct genotype call, where opaque is 100% likelihood). Similarly, LD heatmaps at chromosome 9 around 14 Mb for (D) Unuk-H, (E) Unuk-W, and (F) corresponding heatmap of genotype likelihoods for each sample.

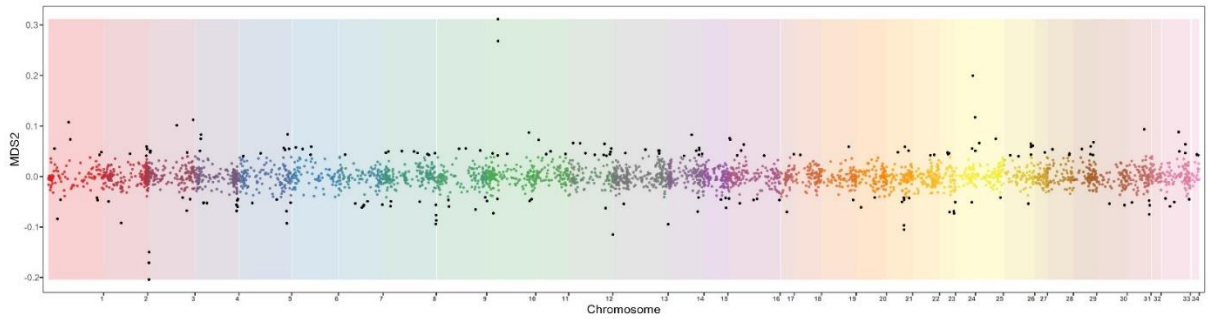


Figure A3: MDS4 plot from *lostruct* output of all sampled individuals, showing outlier regions on chromosomes 3 and 10.

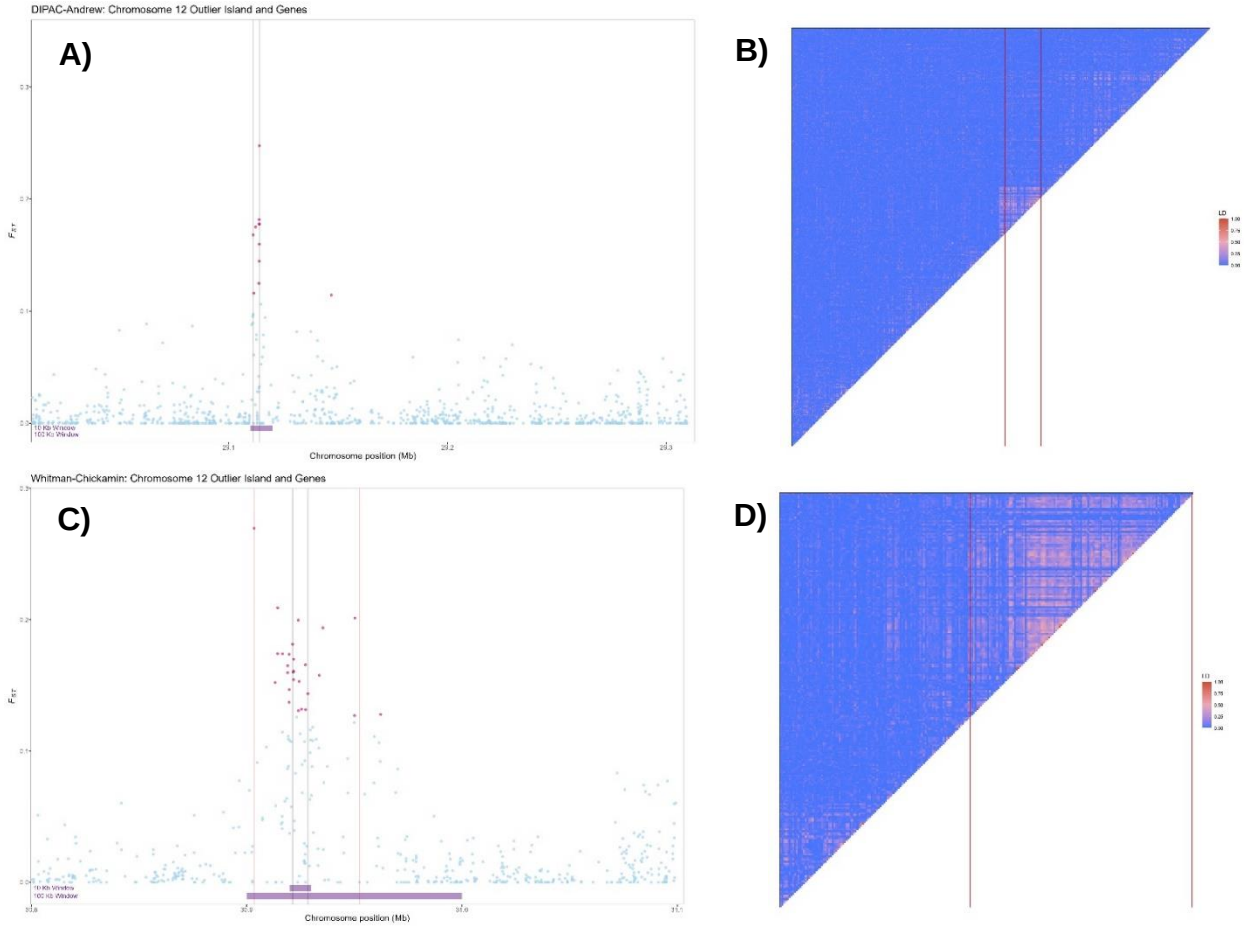


Figure A4: Andrew H/W comparison at chromosome 12 around 29 Mb in the Andrew H/W comparison showing the (A) Outlier F_{ST} peak and (B) corresponding LD heatmap; Outlier F_{ST} peaks on chromosome 12, at approximately 33 Mb in the Chickamin H/W comparison showing the (C) Outlier F_{ST} peak and (D) corresponding LD heatmap. Vertical lines on heatmap represent F_{ST} outlier peak.

Table A1: All identified outlier F_{ST} peaks for all comparisons including the chromosomal position, its relative size, average and maximum F_{ST}, and p-value/significance for the LD tests.

						LD	
Comparison	**Chr**	**Position (Mb)**	**Size (Kb)**	**Average F_{ST}**	**Max F_{ST}**	**Average p-value**	**Significance (%)**
Unuk	2	13.8	35.0	0.043	0.238	<0.0001	100
	3	33.1	49.4	0.058	0.207	0.327	12.3
	4	17.3	4.0	0.095	0.247	<0.0001	100
	4	46.5	84.6	0.065	0.226	<0.0001	100
	4	53.2	6.6	0.048	0.370	<0.0001	100
	5	3.8	46.9	0.032	0.222	<0.0001	100
	6	33.5	37.6	0.051	0.226	<0.0001	100
	6	60.7	7.8	0.085	0.218	0.100	49.2
	6	74.1	6.9	0.102	0.261	<0.0001	100
	9	14.9	60.3	0.075	0.337	<0.0001	100
	10	4.0	14.4	0.032	0.298	<0.0001	100
	14	29.3	6.4	0.108	0.310	<0.0001	100
	18	1.3	25.7	0.044	0.243	<0.0001	100
	28	4.0	79.9	0.038	0.364	0.013	94
	28	6.7	87.9	0.036	0.327	0.557	1.2
	28	40.4	2.1	0.051	0.478	<0.0001	100
	31	25.7	31.1	0.041	0.274	<0.0001	100
Andrew	1	87.4	54.6	0.032	0.221	<0.0001	100
	3	52.7	9.5	0.027	0.185	0.051	72.7
	5	26.5	22.8	0.023	0.156	<0.0001	100
	5	43.8	26.9	0.021	0.151	0.623	0.3
	7	13.7	12.6	0.024	0.177	<0.0001	100
	8	9.5	20.7	0.024	0.339	<0.0001	100
	12	29.1	2.9	0.067	0.248	<0.0001	100
	13	70.9	18.3	0.021	0.143	<0.0001	100
	18	29.3	49.5	0.030	0.212	<0.0001	100
	21	27.0	16.5	0.018	0.242	0.054	72.3
	21	32.8	73.3	0.021	0.201	0.011	94
	21	33.3	36.2	0.015	0.190	0.001	99.7
	25	5.8	67.1	0.020	0.167	<0.0001	100
	25	14.2	85.9	0.032	0.238	<0.0001	100
	33	30.0	2.2	0.051	0.193	<0.0001	100
Chickamin	12	30.9	48.9	0.057	0.270	0.318	12.3
	13	11.3	1.6	0.035	0.274	<0.0001	100
	21	12.3	36.3	0.024	0.171	<0.0001	100
	25	47.2	1.5	0.021	0.212	<0.0001	100
	28	6.5	14.5	0.024	0.224	0.003	99.3

REFERENCES

Ainsworth, C. H., Samhouri, J. F., Busch, D. S., Cheung, W. W. L., Dunne, J., & Okey, T. A. (2011). Potential impacts of climate change on Northeast Pacific marine foodwebs and fisheries. *ICES Journal of Marine Science*, *68*(6), 1217-1229. https://doi.org/10.1093/icesjms/fsr043

Allendorf, F. W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, *5*(2), 181-190. https://doi.org/https://doi.org/10.1002/zoo.1430050212

Álvarez, D., & Nicieza, A. G. (2003). Predator avoidance behaviour in wild and hatchery-reared brown trout: the role of experience and domestication. *Journal of Fish Biology*, *63*(6), 1565-1577. https://doi.org/https://doi.org/10.1111/j.1095-8649.2003.00267.x

Amoroso, R. O., Tillotson, M. D., & Hilborn, R. (2017). Measuring the net biological impact of fisheries enhancement: pink salmon hatcheries can increase yield, but with apparent costs to wild populations. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(8), 1233-1242. https://doi.org/10.1139/cjfas-2016-0334

Arendt, M., Cairns, K. M., Ballard, J. W. O., Savolainen, P., & Axelsson, E. (2016). Diet adaptation in dog reflects spread of prehistoric agriculture. *Heredity*, *117*(5), 301-306. https://doi.org/10.1038/hdy.2016.48

Ayllon, F., Kjærner-Semb, E., Furmanek, T., Wennevik, V., Solberg, M. F., Dahle, G., Taranger, G. L., Glover, K. A., Almén, M. S., Rubin, C. J., Edvardsen, R. B., & Wargelius, A. (2015). The vgll3 Locus Controls Age at Maturity in Wild and Domesticated Atlantic Salmon (Salmo salar L.) Males. *PLOS Genetics*, *11*(11), e1005628. https://doi.org/10.1371/journal.pgen.1005628

Balon, E. K. (2004). About the oldest domesticates among fishes. *Journal of Fish Biology*, *65*(s1), 1-27. https://doi.org/https://doi.org/10.1111/j.0022-1112.2004.00563.x

Barson, N. J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G. H., Fiske, P., Jacq, C., Jensen, A. J., Johnston, S. E., Karlsson, S., Kent, M., Moen, T., Niemelä, E., Nome, T., Næsje, T. F., Orell, P., Romakkaniemi, A., Sægrov, H., Urdal, K., . . . Primmer, C. R. (2015). Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature*, *528*(7582), 405-408. https://doi.org/10.1038/nature16062

Beamish, R. J., Sweeting, R. M., Neville, C. M., Lange, K. L., Beacham, T. D., & Preikshot, D. (2012). Wild chinook salmon survive better than hatchery salmon in a period of poor production. *Environmental biology of fishes*, *94*(1), 135-148. https://doi.org/10.1007/s10641-011-9783-5

Bell, M. A., Aguirre, W. E., & Buck, N. J. (2004). Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution*, *58*(4), 814-824. https://doi.org/10.1111/j.0014-3820.2004.tb00414.x

Besnier, F., Ayllon, F., Skaala, Ø., Solberg, M. F., Fjeldheim, P. T., Anderson, K., Knutar, S., & Glover, K. A. (2022). Introgression of domesticated salmon changes life history and phenology of a wild salmon population. *Evol Appl*, *15*(5), 853-864. https://doi.org/10.1111/eva.13375

Blouin, M. S., Wrey, M. C., Bollmann, S. R., Skaar, J. C., Twibell, R. G., & Fuentes, C. (2021). Offspring of first-generation hatchery steelhead trout (Oncorhynchus mykiss) grow faster in the hatchery than offspring of wild fish, but survive worse in the wild: Possible mechanisms for inadvertent domestication and fitness loss in hatchery salmon. *PLoS One*, *16*(12), e0257407. https://doi.org/10.1371/journal.pone.0257407

Bomblies, K., & Peichel, C. L. (2022). Genetics of adaptation. *Proceedings of the National Academy of Sciences*, *119*(30), e2122152119. https://doi.org/doi:10.1073/pnas.2122152119

Bonamour, S., Chevin, L.-M., Charmantier, A., & Teplitsky, C. (2019). Phenotypic plasticity in response to climate change: the importance of cue variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1768), 20180178. https://doi.org/doi:10.1098/rstb.2018.0178

Bourret, V., O'Reilly, P. T., Carr, J. W., Berg, P. R., & Bernatchez, L. (2011). Temporal change in genetic integrity suggests loss of local adaptation in a wild Atlantic salmon (Salmo salar) population following introgression by farmed escapees. *Heredity (Edinb)*, *106*(3), 500-510. https://doi.org/10.1038/hdy.2010.165

Bradbury, I. R., Lehnert, S. J., Kess, T., Van Wyngaarden, M., Duffy, S., Messmer, A., Wringe, B., Karoliussen, S., Dempson, J. B., Fleming, I. A., Solberg, M. F., Glover, K. A., & Bentzen, P. (2022). Genomic evidence of recent European introgression into North American farmed and wild Atlantic Salmon. *Evolutionary Applications*, *n/a*(n/a). https://doi.org/https://doi.org/10.1111/eva.13454

Christie, M. R., Ford, M. J., & Blouin, M. S. (2014). On the reproductive success of early-generation hatchery fish in the wild. *Evol Appl*, *7*(8), 883-896. https://doi.org/10.1111/eva.12183

Christie, M. R., Marine, M. L., Fox, S. E., French, R. A., & Blouin, M. S. (2016). A single generation of domestication heritably alters the expression of hundreds of genes. *Nature Communications*, *7*(1), 10676. https://doi.org/10.1038/ncomms10676

Christie, M. R., Marine, M. L., French, R. A., Waples, R. S., & Blouin, M. S. (2012). Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. *Heredity*, *109*(4), 254-260. https://doi.org/10.1038/hdy.2012.39

Clucas, G. V., Lou, R. N., Therkildsen, N. O., & Kovach, A. I. (2019). Novel signals of adaptive genetic variation in northwestern Atlantic cod revealed by whole-genome sequencing. *Evolutionary Applications*, *12*(10), 1971-1987. https://doi.org/https://doi.org/10.1111/eva.12861

Crozier, L. G., Burke, B. J., Chasco, B. E., Widener, D. L., & Zabel, R. W. (2021). Climate change threatens Chinook salmon throughout their life cycle. *Commun Biol*, *4*(1), 222. https://doi.org/10.1038/s42003-021-01734-w

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol*, *23*(13), 3133-3157. https://doi.org/10.1111/mec.12796

da Silva Ribeiro, T., Galván, J. A., & Pool, J. E. (2022). Maximum SNP FST Outperforms Full-Window Statistics for Detecting Soft Sweeps in Local Adaptation. *Genome Biology and Evolution*, *14*(10). https://doi.org/10.1093/gbe/evac143

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol Ecol Resour*, *14*(1), 209-214. https://doi.org/10.1111/1755-0998.12157

Doublet, A.-C., Croiseau, P., Fritz, S., Michenet, A., Hozé, C., Danchin-Burge, C., Laloë, D., & Restoux, G. (2019). The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds. *Genetics Selection Evolution*, *51*(1), 52. https://doi.org/10.1186/s12711-019-0495-1

Fang, M., Larson, G., Soares Ribeiro, H., Li, N., & Andersson, L. (2009). Contrasting Mode of Evolution at a Coat Color Locus in Wild and Domestic Pigs. *PLOS Genetics*, *5*(1), e1000341. https://doi.org/10.1371/journal.pgen.1000341

Fast, D., Bosch, W., Johnston, M., Strom, C., Knudsen, C., Fritts, A., Temple, G., Pearsons, T., Larsen, D., Dittman, A., & May, D. (2015). A Synthesis of Findings from an Integrated Hatchery Program after Three Generations of Spawning in the Natural Environment. *North American Journal of Aquaculture*, *77*. https://doi.org/10.1080/15222055.2015.1024360

Fleming, I. A., & Einum, S. (1997). Experimental tests of genetic divergence of farmed from wild Atlantic salmon due to domestication. *ICES Journal of Marine Science*, *54*(6), 1051-1063. https://doi.org/10.1016/s1054-3139(97)80009-4

Flori, L., Fritz, S., Jaffrézic, F., Boussaha, M., Gut, I., Heath, S., Foulley, J.-L., & Gautier, M. (2009). The genome response to artificial selection: a case study in dairy cattle. *PLoS One*, *4*(8), e6595.

Ford, M. J., Berntson, E. A., Moran, P., & McKinney, G. J. (2023). Genomic divergence of hatchery- and natural-origin Chinook salmon (Oncorhynchus tshawytscha) in two supplemented populations. *Conservation Genetics*, *24*(2), 167-179. https://doi.org/10.1007/s10592-022-01491-1

Ford, M. J., Bertson, E., Moran, P., & McKinney, G. (2022). Genomic divergence of hatchery and natural origin Chinook salmon (*Oncorhynchus tshawytscha*) in two supplemented populations (PREPRINT). https://doi.org/https://doi.org/10.21203/rs.3.rs-1789605/v1

Ford, M. J., Fuss, H., Boelts, B., LaHood, E., Hard, J., & Miller, J. (2006). Changes in run timing and natural smolt production in a naturally spawning coho salmon (Oncorhynchus kisutch) population after 60 years of intensive hatchery supplementation. *Canadian Journal of Fisheries and Aquatic Sciences*, *63*(10), 2343-2355.

Fox, E. A., Wright, A. E., Fumagalli, M., & Vieira, F. G. (2019). ngsLD: evaluating linkage disequilibrium using genotype likelihoods. *Bioinformatics*, *35*(19), 3855-3856. https://doi.org/10.1093/bioinformatics/btz200

Fraik, A. K., McMillan, J. R., Liermann, M., Bennett, T., McHenry, M. L., McKinney, G. J., Wells, A. H., Winans, G., Kelley, J. L., Pess, G. R., & Nichols, K. M. (2021). The Impacts of Dam Construction and Removal on the Genetics of Recovering Steelhead (Oncorhynchus mykiss) Populations across the Elwha River Watershed. *Genes*, *12*(1), 89. https://www.mdpi.com/2073-4425/12/1/89

Franklin, I. R., & Frankham, R. (1998). How large must populations be to retain evolutionary potential? *Animal Conservation forum*, *1*(1), 69-70. https://doi.org/10.1111/j.1469-1795.1998.tb00228.x

Fumagalli, M., Vieira, F. G., Korneliussen, T. S., Linderoth, T., Huerta-Sánchez, E., Albrechtsen, A., & Nielsen, R. (2013). Quantifying population genetic differentiation from next-generation sequencing data. *Genetics*, *195*(3), 979-992. https://doi.org/10.1534/genetics.113.154740

Fumagalli, M., Vieira, F. G., Linderoth, T., & Nielsen, R. (2014). ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinformatics*, *30*(10), 1486-1487. https://doi.org/10.1093/bioinformatics/btu041

Gavery, M. R., Nichols, K. M., Goetz, G. W., Middleton, M. A., & Swanson, P. (2018). Characterization of Genetic and Epigenetic Variation in Sperm and Red Blood Cells from Adult Hatchery and Natural-Origin Steelhead, Oncorhynchus mykiss. *G3 (Bethesda, Md.)*, *8*(11), 3723-3736. https://doi.org/10.1534/g3.118.200458

Gilbert, N. (2012). Pink salmon evolve to migrate earlier in warmer waters. *Nature*. https://doi.org/10.1038/nature.2012.10993

Grant, W. S. (2011). Understanding the adaptive consequences of hatchery-wild interactions in Alaska salmon. *Environmental biology of fishes*, *94*(1), 325-342. https://doi.org/10.1007/s10641-011-9929-5

Gustafson, R. G., Waples, R. S., Myers, J. M., Weitkamp, L. A., Bryant, G. J., Johnson, O. W., & Hard, J. J. (2007). Pacific salmon extinctions: Quantifying lost and remaining diversity. *Conservation Biology*, *21*(4), 1009-1020. https://doi.org/10.1111/j.1523-1739.2007.00693.x

Hagen, I. J., Jensen, A. J., Bolstad, G. H., Diserud, O. H., Hindar, K., Lo, H., & Karlsson, S. (2019). Supplementary stocking selects for domesticated genotypes. *Nature Communications*, *10*(1), 199. https://doi.org/10.1038/s41467-018-08021-z

Han, F., Lamichhaney, S., Grant, B. R., Grant, P. R., Andersson, L., & Webster, M. T. (2017). Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Res*, *27*(6), 1004-1015. https://doi.org/10.1101/gr.212522.116

Henson, S. A., Beaulieu, C., Ilyina, T., John, J. G., Long, M., Séférian, R., Tjiputra, J., & Sarmiento, J. L. (2017). Rapid emergence of climate change in environmental drivers of marine ecosystems. *Nature Communications*, *8*(1), 14682. https://doi.org/10.1038/ncomms14682

Hermisson, J., & Pennings, P. S. (2005). Soft Sweeps: Molecular Population Genetics of Adaptation From Standing Genetic Variation. *Genetics*, *169*(4), 2335-2352. https://doi.org/10.1534/genetics.104.036947

Hillestad, B., Makvandi-Nejad, S., Krasnov, A., & Moghadam, H. K. (2020). Identification of genetic loci associated with higher resistance to pancreas disease (PD) in Atlantic salmon (Salmo salar L.). *BMC Genomics*, *21*(1), 388. https://doi.org/10.1186/s12864-020-06788-4

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479-485. https://doi.org/10.1038/nature09670

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A., & Cresko, W. A. (2010). Population Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *PLOS Genetics*, *6*(2), e1000862. https://doi.org/10.1371/journal.pgen.1000862

Huang, K., Andrew, R. L., Owens, G. L., Ostevik, K. L., & Rieseberg, L. H. (2020). Multiple chromosomal inversions contribute to adaptive divergence of a dune sunflower ecotype. *Molecular Ecology*, *29*(14), 2535-2549. https://doi.org/https://doi.org/10.1111/mec.15428

Innan, H., & Kim, Y. (2004). Pattern of polymorphism after strong artificial selection in a domestication event. *Proceedings of the National Academy of Sciences*, *101*(29), 10667-10672. https://doi.org/doi:10.1073/pnas.0401720101

Jin, Y., Olsen, R. E., Harvey, T. N., Østensen, M.-A., Li, K., Santi, N., Vadstein, O., Bones, A. M., Vik, J. O., Sandve, S. R., & Olsen, Y. (2020). Comparative transcriptomics reveals domestication-associated features of Atlantic salmon lipid metabolism. *Molecular Ecology*, *29*(10), 1860-1872. https://doi.org/https://doi.org/10.1111/mec.15446

Jones, L. A., Schoen, E. R., Shaftel, R., Cunningham, C. J., Mauger, S., Rinella, D. J., & St. Saviour, A. (2020). Watershed-scale climate influences productivity of Chinook salmon populations across southcentral Alaska. *Global Change Biology*, *26*(9), 4919-4936. https://doi.org/https://doi.org/10.1111/gcb.15155

Karlsson, S., & Moen, T. (2010). The power to detect artificial selection acting on single loci in recently domesticated species. *BMC Research Notes*, *3*(1), 232. https://doi.org/10.1186/1756-0500-3-232

Karlsson, S., Moen, T., Lien, S., Glover, K. A., & Hindar, K. (2011). Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP-chip. *Molecular Ecology Resources*, *11*(s1), 247-253. https://doi.org/https://doi.org/10.1111/j.1755-0998.2010.02959.x

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Larsen, D. A., Harstad, D. L., Fuhrman, A. E., Knudsen, C. M., Schroder, S. L., Bosch, W. J., Galbreath, P. F., Fast, D. E., & Beckman, B. R. (2019). Maintaining a wild phenotype in a conservation hatchery program for Chinook salmon: The effect of managed breeding on early male maturation. *PLoS One*, *14*(5), Article e0216168. https://doi.org/10.1371/journal.pone.0216168

Le Luyer, J., Laporte, M., Beacham, T. D., Kaukinen, K. H., Withler, R. E., Leong, J. S., Rondeau, E. B., Koop, B. F., & Bernatchez, L. (2017). Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proceedings of the National Academy of Sciences*, *114*(49), 12964-12969. https://doi.org/10.1073/pnas.1711229114

Leitwein, M., Laporte, M., Le Luyer, J., Mohns, K., Normandeau, E., Withler, R., & Bernatchez, L. (2021). Epigenomic modifications induced by hatchery rearing persist in germ line cells of adult salmon after their oceanic migration. *Evolutionary Applications*, *14*(10), 2402-2413. https://doi.org/https://doi.org/10.1111/eva.13235

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754-1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., & Ralph, P. (2018). Local PCA Shows How the Effect of Population Structure Differs Along the Genome. *Genetics*, *211*(1), 289-304. https://doi.org/10.1534/genetics.118.301747

Liu, H., Sørensen, A. C., Meuwissen, T. H. E., & Berg, P. (2014). Allele frequency changes due to hitch-hiking in genomic selection programs. *Genetics Selection Evolution*, *46*(1), 8. https://doi.org/10.1186/1297-9686-46-8

López, M. E., Benestan, L., Moore, J.-S., Perrier, C., Gilbey, J., Di Genova, A., Maass, A., Diaz, D., Lhorente, J.-P., Correa, K., Neira, R., Bernatchez, L., & Yáñez, J. M. (2019). Comparing genomic signatures of domestication in two Atlantic salmon (Salmo salar L.) populations with different geographical origins. *Evolutionary Applications*, *12*(1), 137-156. https://doi.org/https://doi.org/10.1111/eva.12689

Lou, R. N., Jacobs, A., Wilder, A. P., & Therkildsen, N. O. (2021). A beginner's guide to low-coverage whole genome sequencing for population genomics. *Molecular Ecology*. https://doi.org/10.1111/mec.16077

Mäkinen, H., Vasemägi, A., McGinnity, P., Cross, T. F., & Primmer, C. R. (2015). Population genomic analyses of early-phase Atlantic Salmon (Salmo salar) domestication/captive breeding. *Evolutionary Applications*, *8*(1), 93-107. https://doi.org/https://doi.org/10.1111/eva.12230

Malhi, Y., Franklin, J., Seddon, N., Solan, M., Turner, M. G., Field, C. B., & Knowlton, N. (2020). Climate change and ecosystems: threats, opportunities and solutions.

Philosophical Transactions of the Royal Society B: Biological Sciences, *375*(1794), 20190104. https://doi.org/doi:10.1098/rstb.2019.0104

Martinez, V., Dettleff, P. J., Galarce, N., Bravo, C., Dorner, J., Iwamoto, R. N., & Naish, K. (2022). Estimates of Effective Population Size in Commercial and Hatchery Strains of Coho Salmon (Oncorhynchus kisutch (Walbaum, 1792)). *Animals (Basel)*, *12*(5). https://doi.org/10.3390/ani12050647

Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics*, *210*(2), 719-731. https://doi.org/10.1534/genetics.118.301336

Meredith, B. L., Frost, N. D., Reppert, K. S., & Hagerman, G. T. (2022). *Unuk and Chickamin Chinook Salmon Stock Status and Action Plan* (Alaska Department of Fish and Game, Division of Commercial Fisheries, Regional Information Report No. 1J22-13, Issue.

Metcalfe, N. B., Valdimarsson, S. K., & Morgan, I. J. (2003). The relative roles of domestication, rearing environment, prior residence and body size in deciding territorial contests between hatchery and wild juvenile salmon. *Journal of Applied Ecology*, *40*(3), 535-544. https://doi.org/https://doi.org/10.1046/j.1365-2664.2003.00815.x

Mignon-Grasteau, S., Boissy, A., Bouix, J., J., Faure, J. M., Fisher, A. D., Hinch, G. N., Jensen, P., Le Neindre, P., Mormède, P., Prunet, P., Vandeputte, M., & Beaumont, C., C. (2005). Genetics of adaptation and domestication in livestock. *Livestock Production Science*, *93*, 3-14. https://doi.org/10.1016/j.livprodsci.2004.11.001

Muñoz, N. J., Farrell, A. P., Heath, J. W., & Neff, B. D. (2015). Adaptive potential of a Pacific salmon challenged by climate change. *Nature Climate Change*, *5*(2), 163-166. https://doi.org/10.1038/nclimate2473

Naish, K., Seamons, T., Dauer, M., Hauser, L., & Quinn, T. (2013). Relationship between effective population size, inbreeding and adult fitness-related traits in a steelhead (Oncorhynchus mykiss) population released in the wild. *Molecular Ecology*, *22*. https://doi.org/10.1111/mec.12185

Naval-Sanchez, M., McWilliam, S., Evans, B., Yáñez, J. M., Houston, R. D., & Kijas, J. W. (2020). Changed Patterns of Genomic Variation Following Recent Domestication: Selection Sweeps in Farmed Atlantic Salmon [Original Research]. *Frontiers in Genetics*, *11*(264). https://doi.org/10.3389/fgene.2020.00264

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, *29*(1), 1-10. https://doi.org/10.2307/2407137

Neuswanger, J. R., Wipfli, M. S., Evenson, M. J., Hughes, N. F., Rosenberger, A. E., & Jonsson, B. (2015). Low productivity of Chinook salmon strongly correlates with high summer stream discharge in two Alaskan rivers in the Yukon drainage. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*(8), 1125-1137. https://doi.org/10.1139/cjfas-2014-0498

O'Sullivan, R. J., Aykanat, T., Johnston, S. E., Rogan, G., Poole, R., Prodöhl, P. A., de Eyto, E., Primmer, C. R., McGinnity, P., & Reed, T. E. (2020). Captive-bred Atlantic salmon released into the wild have fewer offspring than wild-bred fish and decrease population productivity. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1937), 20201671. https://doi.org/doi:10.1098/rspb.2020.1671

Ohta, T., & Kimura, M. (1969). Linkage disequilibrium due to random genetic drift. *Genetics Research*, *13*(1), 47-55. https://doi.org/10.1017/S001667230000272X

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. *Bioinformatics*, *28*(19), 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, *81*(3), 559-575. https://doi.org/https://doi.org/10.1086/519795

Ravinet, M., Yoshida, K., Shigenobu, S., Toyoda, A., Fujiyama, A., & Kitano, J. (2018). The genomic landscape at a late stage of stickleback speciation: High genomic divergence interspersed by small localized regions of introgression. *PLOS Genetics*, *14*(5), e1007358. https://doi.org/10.1371/journal.pgen.1007358

Raymond, M., & Rousset, F. (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. . *Heredity*, *86:248-249*.

Rennison, D. J., Stuart, Y. E., Bolnick, D. I., & Peichel, C. L. (2019). Ecological factors and morphological traits are associated with repeated genomic differentiation between lake and stream stickleback. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1777), 20180241. https://doi.org/doi:10.1098/rstb.2018.0241

Roberge, C., Einum, S., Guderley, H., & Bernatchez, L. (2006). Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. *Mol Ecol*, *15*(1), 9-20. https://doi.org/10.1111/j.1365-294X.2005.02807.x

Rudman, S. M., Greenblum, S. I., Rajpurohit, S., Betancourt, N. J., Hanna, J., Tilk, S., Yokoyama, T., Petrov, D. A., & Schmidt, P. (2022). Direct observation of adaptive tracking on ecological time scales in Drosophila. *Science*, *375*(6586), eabj7484. https://doi.org/doi:10.1126/science.abj7484

Salomone, P. G., Courtney, K. R., Hagerman, G. T., Fowler, P. A., & Richards, P. J. (2022). *Stikine River and Andrew Creek Chinook Salmon Stock Status and Action Plan* (Alaska Department of Fish and Game, Divisions of Sport Fish and Commercial Fisheries, Regional Information Report No. 1J22-15, Issue.

Slatkin, M. (2008). Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*, *9*(6), 477-485. https://doi.org/10.1038/nrg2361

Smith, C., French, R., Lovtang, J., & Hand, D. (2014). Genetic Composition of the Warm Springs River Chinook Salmon Population Maintained following Eight Generations of Hatchery Production. *Transactions of the American Fisheries Society*, *143*. https://doi.org/10.1080/00028487.2014.931303

Snyder, N. F. R., Derrickson, S. R., Beissinger, S. R., Wiley, J. W., Smith, T. B., Toone, W. D., & Miller, B. (1996). Limitations of Captive Breeding in Endangered Species Recovery. *Conservation Biology*, *10*(2), 338-348. http://www.jstor.org/stable/2386850

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics*, *105*(2), 437-460. https://doi.org/10.1093/genetics/105.2.437

Templin, W. D. (2001). *The History of Propagation and Transportation of Chinook Salmon Oncorhynchus tshawytscha*

Stocks at Hatcheries in Southeast Alaska, 1972–1998. (Alaska Department of Fish and Game, Commercial Fisheries Division, Regional Information Report No. 5J01-05, Issue.

Thériault, V., Moyer, G. R., Jackson, L. S., Blouin, M. S., & Banks, M. A. (2011). Reduced reproductive success of hatchery coho salmon in the wild: insights into most likely

mechanisms. *Molecular Ecology*, *20*(9), 1860-1869. https://doi.org/https://doi.org/10.1111/j.1365-294X.2011.05058.x

Thompson, T. Q., Bellinger, M. R., O'Rourke, S. M., Prince, D. J., Stevenson, A. E., Rodrigues, A. T., Sloat, M. R., Speller, C. F., Yang, D. Y., Butler, V. L., Banks, M. A., & Miller, M. R. (2019). Anthropogenic habitat alteration leads to rapid loss of adaptive variation and restoration potential in wild salmon populations. *Proceedings of the National Academy of Sciences*, *116*(1), 177-186. https://doi.org/doi:10.1073/pnas.1811559115

Tillotson, M. D., Barnett, H. K., Bhuthimethee, M., Koehler, M. E., & Quinn, T. P. (2019). Artificial selection on reproductive timing in hatchery salmon drives a phenological shift and potential maladaptation to climate change. *Evol Appl*, *12*(7), 1344-1359. https://doi.org/10.1111/eva.12730

Utter, F. (1998). Genetic Problems of Hatchery-Reared Progeny Released into the Wild, and how to Deal with them. *Bulletin of Marine Science*, *62*, 623-640.

Van't Hof, A. E., Edmonds, N., Dalíková, M., Marec, F., & Saccheri, I. J. (2011). Industrial Melanism in British Peppered Moths Has a Singular and Recent Mutational Origin. *Science*, *332*(6032), 958-960. https://doi.org/doi:10.1126/science.1203043

Vasemägi, A., Kahar, S., & Ozerov, M. Y. (2016). Genes that affect Atlantic salmon growth in hatchery do not have the same effect in the wild. *Functional Ecology*, *30*(10), 1687-1695. https://doi.org/https://doi.org/10.1111/1365-2435.12635

Vasemägi, A., Nilsson, J., McGinnity, P., Cross, T., O'Reilly, P., Glebe, B., Peng, B., Berg, P. R., & Primmer, C. R. (2012). Screen for Footprints of Selection during Domestication/Captive Breeding of Atlantic Salmon. *Comparative and Functional Genomics*, *2012*, 628204. https://doi.org/10.1155/2012/628204

Waples, R. S., Winans, G. A., Utter, F. M., & Mahnken, C. V. W. (1990). Genetic monitoring of Pacific salmon hatcheries.

Waters, C. D., Hard, J. J., Brieuc, M. S., Fast, D. E., Warheit, K. I., Waples, R. S., Knudsen, C. M., Bosch, W. J., & Naish, K. A. (2015). Effectiveness of managed gene flow in reducing genetic divergence associated with captive breeding. *Evol Appl*, *8*(10), 956-971. https://doi.org/10.1111/eva.12331

Waters, C. D., Hard, J. J., Brieuc, M. S. O., Fast, D. E., Warheit, K. I., Knudsen, C. M., Bosch, W. J., & Naish, K. A. (2018). Genomewide association analyses of fitness traits in captive-reared Chinook salmon: Applications in evaluating conservation strategies. *Evolutionary Applications*, *11*(6), 853-868. https://doi.org/https://doi.org/10.1111/eva.12599

Waters, C. D., Hard, J. J., Fast, D. E., Knudsen, C. M., Bosch, W. J., & Naish, K. A. (2020). Genomic and phenotypic effects of inbreeding across two different hatchery management regimes in Chinook salmon. *Molecular Ecology*, *29*(4), 658-672. https://doi.org/https://doi.org/10.1111/mec.15356

Yáñez, J. M., Naswa, S., López, M. E., Bassini, L., Correa, K., Gilbey, J., Bernatchez, L., Norris, A., Neira, R., Lhorente, J. P., Schnable, P. S., Newman, S., Mileham, A., Deeb, N., Di Genova, A., & Maass, A. (2016). Genomewide single nucleotide polymorphism discovery in Atlantic salmon (Salmo salar): validation in wild and farmed American and European populations. *Molecular Ecology Resources*, *16*(4), 1002-1011. https://doi.org/https://doi.org/10.1111/1755-0998.12503

VITA

PERSONAL BACKGROUND

Natasha Howe, born 25 May, 1994 in Torrance, California to Brett Howe and Heidi Simone.

Graduated from Mission Viejo High School in Mission Viejo, California in June of 2012.

EDUCATION

M.S. Biology, Texas Christian University

M.B.A., Seattle University

B.S. Environmental Engineering, Seattle University

AWARDS AND FUNDING

Science & Engineering Research Center (SERC) Grant – Texas Christian University - *$2000*

Adkins Fellowship – Texas Christian University - *$3600*

TEACHING EXPERIENCE

Teaching Assistantship, Texas Christian University

- Introductory Biology I & II Laboratory – Spring 2021 and Spring 2023
- Contemporary Issues in Biology Laboratory – Fall 2021 and Fall 2022

ABSTRACT

GENOMIC EVIDENCE FOR HATHERY-INDUCED DOMESTICATION SELECTION IN CHINOOK SALMON, *ONCORHYNCHUS TSHAWYTSCHA*

By Natasha Howe, M.S. 2023

Department of Biology

Texas Christian University

Thesis Advisor: Dr. Matthew C. Hale, Associate Professor of Biology

Salmon hatcheries are widely used across the Pacific, yet hatchery fish generally have reduced fitness compared to their wild counterparts. Domestication selection, or adaptation to the hatchery environment, poses a risk to wild populations if introgression between hatchery and wild fish occurs. In this study, we examined three separate hatchery populations of Chinook salmon, *Oncorhynchus tshawytscha*, and their corresponding wild progenitor populations using whole genome sequencing. Each hatchery population was compared to its wild progenitor population using multiple metrics of genomic divergence. While evaluating population-level genomic differentiation (F_{ST}), we discovered outlier peaks in each hatchery-wild pair, although no outliers were shared across the comparisons. Further analyses indicated that these relatively small peaks are likely due to genetic hitchhiking on hatchery-selected alleles. Overall, our genome-wide analyses provide fine-scale genetic evidence for domestication and highlight the need to assess if certain management practices can mitigate genetic risks despite multiple pathways of domestication.