WHERE ARE YOU GOING? USING GENETIC METHODS TO DETERMINE THE IMPORTANCE OF SEVERAL GENES IN LIFE-HISTORY DEVELOPMENT OF RAINBOW TROUT (*ONCHORHYNCHUS MYKISS*).

by

Anna Claire Pilkington

Submitted in partial fulfillment of the
requirements for Departmental Honors in
the Department of Biology
Texas Christian University
Fort Worth, Texas

May 4, 2020

WHERE ARE YOU GOING? USING GENETIC METHODS TO DETERMINE THE IMPORTANCE OF SEVERAL GENES IN LIFE-HISTORY DEVELOPMENT IN RAINBOW TROUT (*ONCHORHYNCHUS MYKISS*).

Project Approved:

Supervising Professor: Matt Hale, Ph.D.

Department of Biology

John Horner, Ph.D.

Department of Biology

Gina Hill, Ph.D. RD, LD

Department of Nutritional Sciences

ABSTRACT

Migration is broadly defined as the movement of individuals from one area to another to take advantage of seasonal resources. Rainbow trout (*Oncorhynchus mykiss*) exist as two different types (commonly known as rainbow and steelhead trout) that exhibit different migratory patterns. While the decision to migrate or stay resident is complex, it is well-known that there is a relationship between the migratory patterns of parents and the migratory patterns of their offspring, suggesting a strong heritable basis. Although several genes have been identified that are associated with migration in rainbow trout, how widespread these polymorphisms are between different populations of trout remains unknown. This project aimed to determine polymorphisms that may be contributing to the migratory versus stationary phenotype in two populations from Southeast Alaska. Sequencing data collected from two loci, Omy16 and Omy17, led to the identification of two SNPs with significantly different allelic frequencies between residents and migrants, suggesting that allelic differences exist between Sashin Lake and Sashin Creek. It is possible that these alleles can be used to determine where migrating smolts originated from. Future studies will be conducted on larger sample sizes and different populations to determine if these loci of interest can be used across populations and in conservation efforts.

Introduction.

Migration is broadly defined as the movement of individuals to take advantage of seasonal resources (Liedvogel et al. 2011). Migration most commonly occurs to serve the purpose of resource attainment, in the form of food or shelter. However, several migratory species are declining at rates disproportional to their resident counterparts and conservation biologists are trying to understand the reasons for this disparity. Rainbow trout, *Oncorhynchus mykiss*, exist as two different ectomorphs; resident and migrants. Resident forms are referred to as rainbow trout, whereas migrants are commonly known as steelhead trout. Migratory populations are declining at rates not mirrored in resident populations and therefore provide an excellent model system for conservation genetics research to more fully understand the reasons for these declines (Liedvogel et al. 2011, FWS 2011).

Rainbow trout spend their lives in freshwater whereas the steelhead trout spend the first portion of their life in freshwater, followed by a migration to the ocean, and then returning to their natal streams to spawn (Hu et al. 2014; Narum et al. 2018). To prepare for the salt water environment, steelhead must undergo a series of physiological, behavioral, and morphological changes known as smoltification (McCormick et al. 2003; Hoar 1988). These changes are large-scale and include increased salinity tolerance, increased body length, decreased lipid composition, and acquiring a silvery color to blend in with the pelagic environment (McCormick et al. 2003; Björnsson et al. 2011). Although environmental cues, such as increased day-length and temperature, are known to regulate this process, the propensity to develop into a migratory steelhead or a resident rainbow trout appears to be heritable, suggesting underlying genetic effects (Hoar 1976; Thrower et al. 2004; Hecht et al. 2012; Hu et al. 2014).

Although previous studies have suggested that migratory behavior appears to be heritable, few studies have been able to determine the genetic basis of this behavior due to the impracticalities of studying migratory animals in a controlled environment (Liedvogel et al. 2011). The rainbow trout provide an excellent opportunity to assess these migratory differences because a) the two ecotypes – i.e., rainbows and steelhead - are easily distinguishable from one another, and b) it is one of the few migratory animals to have its genome published (Pearse et al. 2019). In determining the genetic basis of the two ecotypes, rainbow and steelhead trout from Sashin Creek in southeastern Alaska have become a model population. This system contains two distinct populations, a mostly resident population that was founded by steelhead trout that were moved above two barrier waterfalls, and a migratory population that breeds below the barrier waterfalls (Thrower et al. 2004; Figure 1). Previous studies have used this system to evaluate and identify genetic differences between migratory and resident populations, but given that both migratory and resident organisms can produce migratory offspring, it is important to develop genetic markers that are indicative of parental origin.

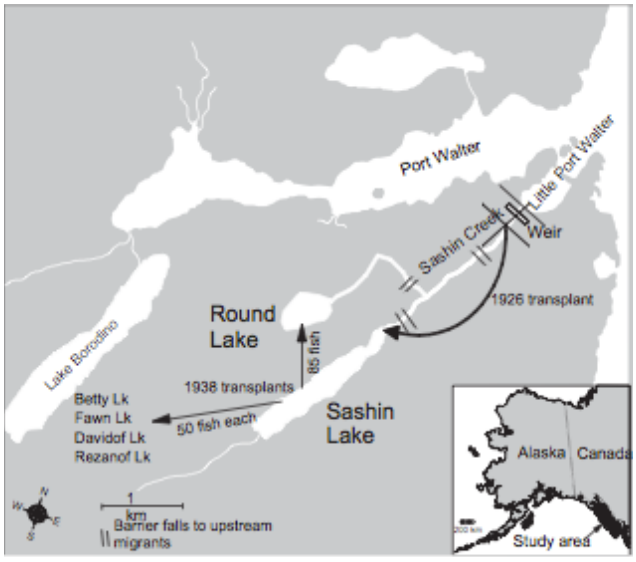

Figure 1. Map of Sashin Creek and the location of the migrant and resident rainbow trout populations.

The development of genetic markers that associate with life history will aid in being able to determine the ultimate ecotype development of rainbow trout before they begin smoltification and downstream migration. This is of growing importance because eleven steelhead trout populations in the pacific northwest United States are in decline, with many of them already extinct (Gustafson et al. 2007). To aid in the recovery of the migratory form, resident rainbow trout populations that have become sequestered in lakes could be used to recover migratory steelhead trout populations (Thrower et al. 2008). However, this restoration is not possible without the ability to determine if, and to what extent, migratory trout are being produced from resident parents. The single nucleotide polymorphisms (SNPs) that are unique to each ecotype can function as both a genetic marker and provide insight into the genetic control of migration in other organisms (Nichols et al. 2008; Hecht et al. 2012; Hale et al. 2013).

The goals of this project were twofold: 1) to identify polymorphic positions that are consistently different between migrant and resident rainbow trout from Sashin Creek, Alaska, and 2) to use the polymorphisms to accurately determine if a migrant leaving Sashin Creek to begin its migratory life-cycle had resident or migrant parents. In doing so, we hope to identify the genes important in migration and provide candidate genes that could be useful in other populations of rainbow trout with decreasing numbers of migratory steelhead.

Methods.

In this project, a forward genetics approach was used to identify polymorphisms that could accurately determine if a migratory trout leaving Sashin Creek had resident or migrant parents. Primers were created to be used in polymerase chain reaction (PCR) to amplify the genes of

interest. The sequence of the gene of interest was determined by genetic analyzer. These sequences were further analyzed to create contigs that allow for comparisons to be made between specific base pairs of the migrants and the residents. This process was repeated for multiple regions of the rainbow trout genome in an attempt to identify SNPs associated with the development of the two ecotypes.

Samples

Samples utilized in this study were generated through collection of migratory and resident adult rainbow fish from natural populations in Sashin Creek and Sashin Lake, Alaska, respectively. Collection of samples occurred from May 2010 and then yearly until May 2016. All fish were non-lethally sampled. Migratory fish were sampled by collection at a weir that stops returning from steelhead from entering Sashin Creek to spawn. Fish were removed from the weir, fin clipped for genetic analysis, weighed and measured, and then released into Sashin Creek to complete their life cycle. Residents were sampled by using hook-and-wire traps in Sashin Lake. Residents were weighed and measured, and fin clips were taken for genetic analysis. Fin clips were placed in 70% ethanol and stored in at -20 °C.

DNA Extraction

DNA was extracted from fin clips using the Qiagen DNeasy Blood and Tissue DNA extraction kit (following manufacturer's protocol; Qiagen, Hilden, Germany) and diluted to a standard concentration of 50 ng/μL. Because of the known sex bias in life history, fish that had not reached sexual maturity were sexed with *OmyY1*, a marker on the male-specific portion of the Y-chromosome (Brunelli et al. 2008). The sexing PCR protocol followed Brunelli et al. 2008,

except annealing conditions were 60°C for 50 seconds. Sex was determined by running 4 µL of PCR product on a 1.5% agarose gel stained with Gel Red and viewed under UV light.

Primer Production

Primers were created for amplification of a total of thirteen different loci that were in regions of the genome previously thought to be important in life history development in rainbow trout from Sashin Creek (Hale et al. 2013; Weinstein et al. 2019). Primers were designed in Primer3 using default parameters.

PCR

Viability of each primer was tested through PCR on both resident and migrant rainbow trout and five of these primer pairs were further sequenced (see table 1).Each PCR reaction was run on a Bio-Rad T100 Thermal Cycler. The thermal cycling process consisted of a holding stage at 95 °C for 10 minutes, a cycling stage of 35 or 40 cycles of denaturing at 95 °C for 30 seconds, annealing at primer-dependent temperature for 30 seconds (see table 1), and extension for 30 seconds at 72 °C. After the final cycle, there was a 5 minute hold at 72 °C. The annealing temperature and number of cycles varied depending on the primer used. Negative controls, i.e., nuclease free water, were used for all PCRs to check for non-specific amplification and DNA contamination.

Reactions were carried out using one of two master mixes. The first, denoted as Taq in Table 1, contained 0.6 µL 25 mmol $MgCl_2$, 0.1 µL Taq polymerase, 2 µL 5x Green GoTaq Flexi Buffer, 1 µL 2 mmol dNTPs , 0.5 µL forward and reverse primer, and 4.3 µL water, along with 1 µL

diluted DNA to equal a total volume of 10 µL. The second master mix, denoted as Accustart in Table 1, contained 5 µL Accustart II PCR SuperMix, 0.5 µL of forward primer, 0.5 µL reverse primer, and 3 µL water, along with 1 µL diluted DNA to equal a total volume of 10 µL. After running the PCR, 5 µL of PCR product were run on an agarose gel to confirm the presence of the desired product, along with ensuring that the product was clean (i.e., only amplified one product with minimum primer-dimer). The agarose gels were made using 1.5 g Agarose II and 100 mL SB buffer. Samples were loaded using GelRed loading dye, and length was estimated via comparison to BioLabs Quick Load 100-bp DNA ladder. All gels were then visualized using UV light.

Loci	Methods	Annealing Temperature (°C)	Number of Cycles
Omy5	Taq	55.7	40
Omy8	Accustart	65	40
Omy16	Accustart	68	35
Omy17	Accustart	53.8	40
Omy28	Accustart*	59	40

Table 1. Methods, annealing temperature, and number of cycles used for the different loci. *Indicates that 0.25 µL of each primer and 3.5 µL water were used in place of the originally described amounts

ExoSap PCR

After confirming the presence of the desired DNA product, the amplified DNA was purified using an ExoSap PCR. For this reaction, 1.19 µL water, 0.2 µL ExoSap, 0.2 µL rSap, and 0.7 µL

Neb2 were combined to create a master mix. From this, 2 µL of master mix were added to 5 µL of PCR product, to create a total volume of 7 µL. This was then run on the thermal cycler for 15 minutes at 37 °C and for 15 minutes at 80 °C.

Big Dye PCR

Sequencing was conducted using BigDye V3.1 chemistry. The Big Dye PCR consisted of 5.75 µL water, 1.0 µL of either forward or reverse primer, 1.75 µL BigDye Terminator 5x Sequencing Buffer, and 0.5 µL BigDye, along with 1 µL of DNA product from the ExoSap PCR, to equal a total volume of 10 µL. On the thermal cycler, there was first a time of 3 minutes at 96 °C followed by 30 cycles of 10 seconds at 96 °C, 5 seconds at 50 °C, and 2.5 minutes at 60 °C.

Reaction Clean-Up

To purify the samples before sequencing, 20 µL of BET was added to the product from the BigDye reaction and mixed thoroughly. After magnetizing the beads for 20 seconds, all the liquid was removed, leaving only the beads behind. The beads are washed twice with 70% ethanol, 100 µL for the first wash and 120 µL for the second. The DNA was eluted from the beads using 40 µL EDTA and mixing. While on the magnet, all the liquid was transferred to a semi-skirted plate, leaving behind the beads. The plate was covered, spun down, and placed in the sequencer.

Sequencing

Samples were sequenced on an Applied Biosystems 3130x Genetic Analyzer, using sequencing results, instrument protocol UltraSeq, and analysis protocol 313POP-7. Upon completion of

sequencing, and after determining the presence of clean peaks using SeqA6, the data was analyzed using Sequencher 5.4.6. Contigs were created to identify SNPs seen between the samples, specifically between migrant and resident rainbow trout.

Results.

PCR optimization was attempted on thirteen loci, of which five were successfully amplified. Each of these loci resided on a different chromosome; Omy5, Omy8, Omy28, Omy16, and Omy17. All of these chromosomal regions have been found to be important in the genetic control of migration in other studies and, thus, were the focus of the rest of this study.

Although PCR was successful with Omy8 and Omy28, it was not possible to obtain usable sequence information, and so these two loci were not included in subsequent analyses. Sequencing of Omy5 was successful for 5 resident and 3 migrants but there were no SNPs found in the sequence that was available for analysis. Because previous studies have shown significant differences amongst migrants and residents at this locus, it is likely that these results were due to the small number of samples sequenced. A larger sample size might have found evidence of SNPs associating with life history.

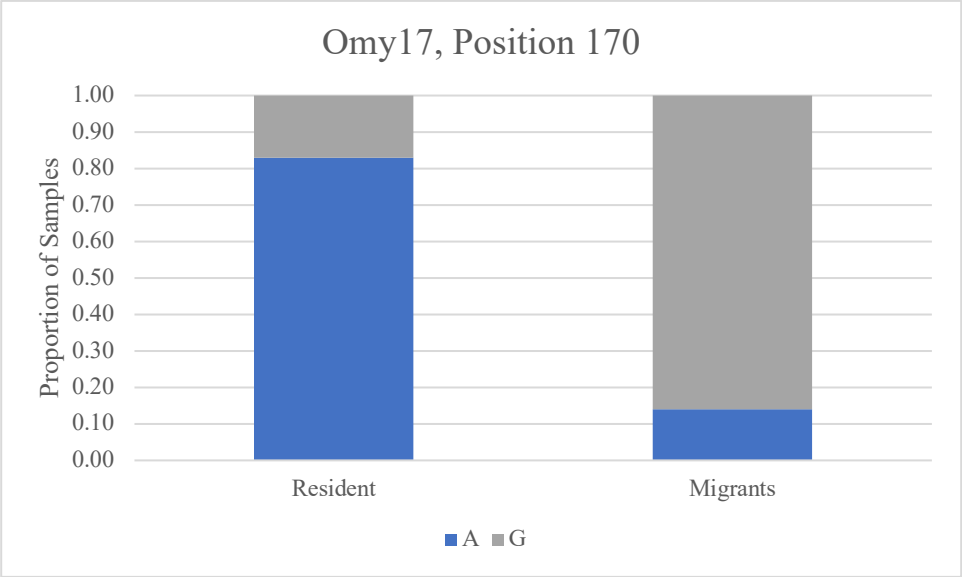

Figure 2. Proportion of samples at position 170 of the Omy17 locus. For residents, n=12 and for migrants, n=7. χ^2=8.92, p<0.05

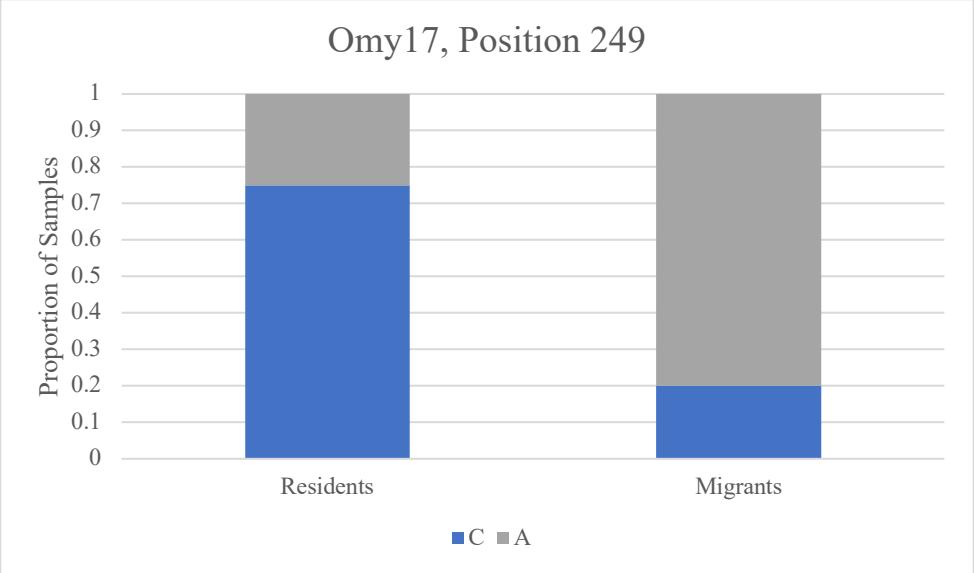

Figure 3. Proportion of samples at position 249 of the Omy17 locus. For residents, n=4 and for migrants, n=6. χ^2=3.66, p<0.05

Sequencing of the Omy17 locus identified fifteen possible SNPs between the migrant and resident rainbow trout, with two of the most compelling shown in figures 2 and 3. At nucleotide

position 170 of this locus, 12 residents and 7 migrants were sequenced and analyzed. There was a significant difference between resident and migratory fish at this position.

At nucleotide position 249 of the Omy17 locus, 4 residents and 6 migrants were sequenced and analyzed. There was no significant difference between migrant and resident rainbow trout.

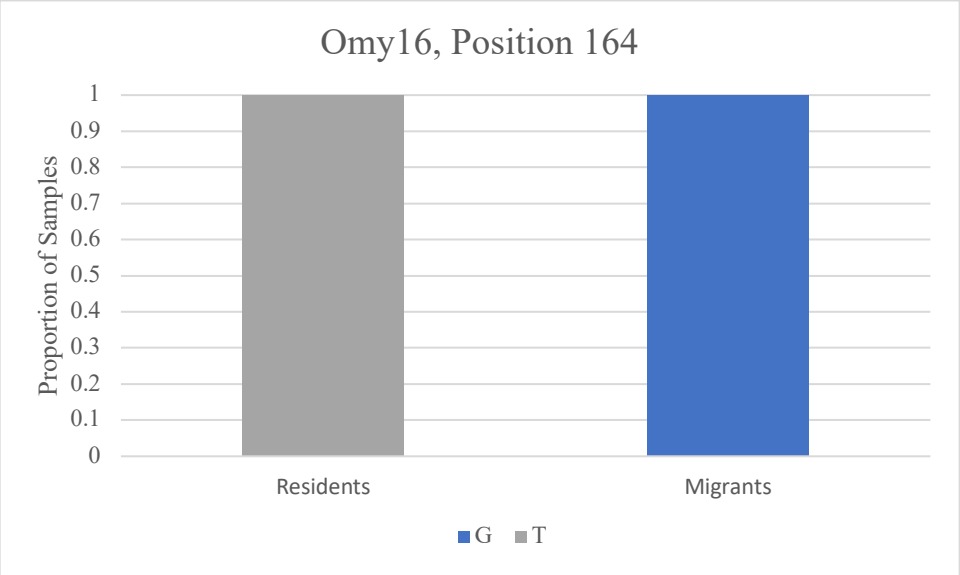

Figure 4. Proportion of samples at position 164 of the Omy16 locus. For residents, n=1 and for migrants, n=4. χ^2=5.00, p<0.05

While sequencing analysis was completed on the Omy16 locus, few samples (1 resident and 4 migrants) produced a clean result that could be subsequently analyzed. Again, there was a significant difference between migrant and resident rainbow trout (Figure 4).

Discussion.

Previous genomic studies have identified regions that may be useful in locating migratory markers in steelhead trout and these regions were focused on in this study. These regions are

spread throughout the genome and have been identified by characterizing linkage groups that contain quantitative trait loci (QTL) associated with migration and migratory-related traits (Nichols et al. 2008; Hale et al. 2013; Weinstein et al. 2019). For example, Omy5 has a role in the growth rate and body morphology observed in the migrants versus the residents (Nichols et al. 2008). Omy5 has consistently been shown to contain a region where several QTL associated with migratory traits localize and has been demonstrated to contain genes that play a role in the regulation of seasonal rhythms, an important trait in determining when to migrate (Martínez et al. 2011; Hecht et al. 2012). In this study, no SNPs were found in the sequencing region of Omy5 and therefore, it could not be further tested. Differences in two genes located on Omy28 have been shown to be associated with timing of migration in both rainbow trout and Chinook salmon (*Oncorhynchus tshawytscha*), suggesting conservation of this difference among related species (Micheletti et al. 2018; Narum et al. 2018). The similarities observed across multiple studies demonstrates the importance of these genomic regions because they are being conserved amongst differing populations, providing validity into their role in migration. In addition, such regions also have the potential to be useful in assessing whether a migrant trout had migratory or resident parents.

Former studies have determined that Omy16 and Omy17 contain regions that code for the physiological change seen in the steelhead trout during smoltification but little has been documented regarding SNPs between steelhead and rainbow trout that could be contributing to these observed phenotypical differences (Nichols et al. 2008; Hecht et al. 2012). In this study, SNPs were found on Omy16 and Omy17 that can possibly be used as markers to determine rainbow trout life history before smoltification commences. Thus, these polymorphisms may be

useful in determining if a resident population can replenish the diminishing steelhead trout population. If a resident trout has the SNPs present that are important for the migratory phenotype, then it is possible that resident will produce migratory offspring at a higher frequency than a resident with the alternative allele (Thrower et al. 2008). Although these findings are exciting, it is important to note that the sample sizes for both loci were small. In order to ensure that these alleles are truly useful in determining origin of migratory smolts, it will be necessary to ground-truth in a larger number of samples.

Along with analysis of more samples, work should also be done to locate SNPs that are found throughout the rainbow trout genome. There are undoubtedly SNPs present at multiple loci throughout that could be used to gain greater understanding regarding the use of resident populations in conservation efforts. Despite the limitations of this study, SNPs on Omy16 and Omy17 were determined and can be utilized in future studies.

The main aims of this study were to identify SNPs that can be utilized to distinguish between resident and migrant rainbow trout and to determine if the SNPs at different loci can be used to pinpoint the phenotype of a migrant trout's parents. Loci that appear to be segregating between phenotypes were identified and it is possible that these loci can be useful in other populations and in conservation efforts. It should be noted that larger sample sizes and testing in both the Sashin system and other populations of interest are still needed to conclusively make a link between genotype and the observed phenotype. In addition, the second aim of screening candidate SNPs in returning steelhead to determine if they originated from the lake or the creek was not met. However, Omy16 and Omy17 might be appropriate in prospective tests used to

determine the origin of returning steelhead. Future studies should focus on the genotyping of larger sample sizes to allow for a more complete understanding of how prevalent these polymorphisms are in the population of interest. With this knowledge, confirmation of alleles as being lake or stream specific will be available while also giving greater insight into the populations at hand.

References.

Björnsson BT, Stefansson SO, McCormick SD. 2011. Environmental endocrinology of salmon smoltification. General and Comparative Endocrinology. 170(2):290–298. doi:10.1016/j.ygcen.2010.07.003. [accessed 2020 Feb 2]. https://linkinghub.elsevier.com/retrieve/pii/S0016648010002418.

Brunelli JP, Wertzler KJ, Sundin K, Thorgaard GH. 2008. Y-specific sequences and polymorphisms in rainbow trout and Chinook salmon. Genome. 51(9):739–748. doi:10.1139/G08-060. [accessed 2020 Feb 28]. https://www.nrcresearchpress.com/doi/10.1139/G08-060.

Gustafson RG, Waples RS, Myers JM, Weitkamp LA, Bryant GJ, Johnson OW, Hard JJ. 2007. Pacific salmon extinctions: quantifying lost and remaining diversity. Conservation Biology. 21(4):1009–1020. doi:10.1111/j.1523-1739.2007.00693.x. [accessed 2020 Feb 2]. https://conbio.onlinelibrary.wiley.com/doi/abs/10.1111/j.1523-1739.2007.00693.x.

Hale MC, Thrower FP, Berntson EA, Miller MR, Nichols KM. 2013. Evaluating adaptive divergence between migratory and nonmigratory ecotypes of a salmonid fish, *Oncorhynchus mykiss*. G3. 3(8):1273–1285. doi:10.1534/g3.113.006817. [accessed 2020 Feb 5]. http://g3journal.org/lookup/doi/10.1534/g3.113.006817.

Hecht BC, Thrower FP, Hale MC, Miller MR, Nichols KM. 2012. Genetic architecture of migration-related traits in rainbow and steelhead trout, Oncorhynchus mykiss. G3: Genes, Genomes, Genetics. 2(9):1113–1127. doi:10.1534/g3.112.003137. [accessed 2020 Feb 5]. https://www.g3journal.org/content/2/9/1113.

Hoar WS. 1976. Smolt transformation: evolution, behavior, and physiology. :20.

Hoar, W.S., 1988. The physiology of smolting salmonids. In: Hoar, W.S., Randall, D.J. (Eds.), Fish Physiology, vol. XIB. Academic Press, NY, pp. 275–343.

Hu G, Wang C, Da Y. 2014. Genomic heritability estimation for the early life-history transition related to propensity to migrate in wild rainbow and steelhead trout populations. Ecology and Evolution. 4(8):1381–1388. doi:10.1002/ece3.1038. [accessed 2020 Feb 5]. https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.1038.

Liedvogel M, Åkesson S, Bensch S. 2011. The genetics of migration on the move. Trends in Ecology & Evolution. 26(11):561–569. doi:10.1016/j.tree.2011.07.009. [accessed 2020 Feb 5]. http://www.sciencedirect.com/science/article/pii/S0169534711002072.

Martínez A, Garza JC, Pearse DE. 2011. A microsatellite genome screen identifies chromosomal regions under differential selection in steelhead and rainbow trout. Transactions of the American Fisheries Society. 140(3):829–842. doi:10.1080/00028487.2011.588094. [accessed 2020 Feb 6]. https://afspubs.onlinelibrary.wiley.com/doi/abs/10.1080/00028487.2011.588094.

McCormick SD, O'Dea MF, Moeckel AM, Björnsson BT. 2003. Endocrine and physiological changes in Atlantic salmon smolts following hatchery release. Aquaculture. 222(1):45–57.

doi:10.1016/S0044-8486(03)00101-7. [accessed 2020 Feb 5]. http://www.sciencedirect.com/science/article/pii/S0044848603001017.

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(1):140. doi:10.1186/s12862-018-1255-5. [accessed 2020 Feb 7]. https://doi.org/10.1186/s12862-018-1255-5.

Narum SR, Di Genova A, Micheletti SJ, Maass A. 2018. Genomic variation underlying complex life-history traits revealed by genome sequencing in Chinook salmon. Proceedings of the Royal Society B: Biological Sciences. 285(1883):20180935. doi:10.1098/rspb.2018.0935. [accessed 2020 Feb 7]. https://royalsocietypublishing.org/doi/full/10.1098/rspb.2018.0935.

Nichols KM, Edo AF, Wheeler PA, Thorgaard GH. 2008. The genetic basis of smoltification-related traits in Oncorhynchus mykiss. Genetics. 179(3):1559–1575. doi:10.1534/genetics.107.084251. [accessed 2020 Feb 2]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2475755/.

Pearse W, Legendre P, Peres-Neto P, Davies J. 2019. The interaction of phylogeny and community structure: Linking the community composition and trait evolution of clades. Global Ecology and Biogeography. 28. doi:10.1111/geb.12938.

Thrower FP, Hard JJ, Joyce JE. 2004. Genetic architecture of growth and early life-history transitions in anadromous and derived freshwater populations of steelhead. Journal of Fish Biology. 65(s1):286–307. doi:10.1111/j.0022-1112.2004.00551.x. [accessed 2020 Feb 2]. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0022-1112.2004.00551.x.

Thrower FP, Joyce JE, Celewycz, Adrian G. M Patrick. 2008. The ootential importance of reservoirs in the western United States for the recovery of endangered populations of anadromous steelheads. American Fisheries Society Symposium. 62:309–324.

Weinstein SY, Thrower FP, Nichols KM, Hale MC. 2019. A large-scale chromosomal inversion is not associated with life history development in rainbow trout from Southeast Alaska. PLOS ONE. 14(9):e0223018. doi:10.1371/journal.pone.0223018. [accessed 2020 Feb 6]. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0223018.