MEASURING ASSOCIATIONS BETWEEN POLYMORPHISMS IN THE *GREB1-L* GENE AND THE DEVELOPMENT OF DIFFERENT MIGRATORY PHENOTYPES IN *ONCORHYNCHUS MYKISS*

by

BEN SHOWALTER

Submitted in partial fulfillment of the

requirements for Departmental Honors in

the Department of Biology

Texas Christian University

Fort Worth, Texas

May 8, 2017

MEASURING ASSOCIATIONS BETWEEN POLYMORPHISMS IN THE *GREB1-L* GENE AND THE DEVELOPMENT OF DIFFERENT MIGRTORY PHENOTYPES IN *ONCORHYNCHUS MYKISS*

Project Approved:

Supervising Professor: Matthew Hale, Ph.D.

Department of Biology

Dean Williams, Ph.D.

Department of Biology

David Minter, Ph.D.

Department of Chemistry

Abstract

Migration can be defined as the movement of individuals to new areas followed by a return to their natal area to breed. *Oncorhynchus mykiss,* otherwise known as rainbow trout, exhibits two life history strategies: resident rainbow trout and migratory steelhead trout. Previous research has shown that the decision to migrate is highly heritable, that there are regions of the genome that are associated with migration, and that there are differentially expressed genes between residents and migrants, all confirming that migration has a genetic component. Recently, interest has focused on the *GREB1-L* gene as studies in several populations of rainbow trout have found alleles within this gene that are associated with migration. This project is aimed to measure allelic associations between *GREB1-L* and migratory life history in rainbow trout from Sashin Creek, Alaska. Sequence data suggest that all individuals, regardless of their migratory trajectory, had the alleles associated with residency. These results confirm that *GREB1-L* is locally adapted to different environments and that there are population-specific genetic effects that determine the migratory life history of rainbow trout.

Introduction

Migration can be defined as the movement of individuals to new areas followed by a return to their natal area to breed. Rainbow trout (*Oncorhynchus mykiss)* are one of many species to perform this phenomenon (Björnsson et al., 2011). This species is comprised of two different migratory life-history strategies: resident rainbow trout and migratory steelhead trout (Thrower et al., 2004). At about two years of age, these fish make the decision to become rainbow or steelhead trout (Hale et al., 2016). For individuals to cope with the demands of migration, migratory trout must change from being adapted to freshwater to being adapted to salt water (Nichols et al., 2008). This process, known as smoltification, affects the morphology, behavior and physiology of these migratory trout (Nichols et al., 2008). Migrants move to the ocean where they spend between two and four years feeding until they return to their natal fresh water streams to reproduce (Nichols et al., 2008). If the resident pathway is chosen, the individual will remain in fresh water for the duration of their lives (Nichols et al., 2008).

The scientific community has long been interested in how migration is inherited in rainbow trout. It is now known that both environmental and genetic factors are involved in the migratory decision (Hale et al., 2016; Laura Liberoff et al., 2015). For example, the closer the fish's freshwater habitats are to the ocean, the more likely they are to become migrants (Laura Liberoff et al., 2015). Evidence from quantitative genetics studies have shown that migration in rainbow trout is highly heritable, as individuals who had migratory parents are more likely to migrate and successfully return to spawn than an individual with resident parents (Thrower et al., 2004; Hecht et al., 2015). This association has caused interest to find the specific genes that may be

involved in the migratory decision. A study in 2008 revealed several regions of the rainbow trout genome associated with variation in several morphological measurements known to be associated with the smoltification process such as skin silvering, changes in body shape and growth rate (Nichols et al., 2008). Similarly, a study published in 2012 found strong associations between smoltification related traits and two chromosomal positions, suggesting that master genetic controls for the migratory decision could be within these regions (Hecht et al., 2012). Two studies published in 2013 found significant single nucleotide polymorphisms (SNPs) using specific regions of the genome associated with migration in two populations of rainbow trout (Hale et al., 2013; Hecht et al., 2013). Although these previous studies prove that migration in rainbow trout has a genetic basis, quantitative trait locus (QTL) methods and genome-wide association studies (GWAS) make it difficult to locate the causative gene. In addition, these studies suggest that the genetic basis for migration varies between populations, as there was little concordance in the location of alleles associated with migration. Other studies have taken a different approach to targeting these genes associated with migration by using RNA sequencing to measure differential expression between migrants and residents (McKinney et al., 2015; Hale et al., 2016). Both studies measured gene expression in the brain and revealed that many hundreds of genes are differentially expressed between residents and migrants.

It is clear that there is a genetic basis to the migratory life-history strategy, but the identification of specific genes associated with this decision has proven to be a more difficult task. Knowing these specific genes would not only be of interest to understanding the evolutionary basis of migration, but would also be of interest in conservation studies (Goetz et al., 2015). Many anadromous populations of Pacific salmon are decreasing, especially in the lower 48 states

(Busby et al., 1996). As a result, conservation biologists as well as fishermen are interested in restocking migratory populations of salmonids. An increased understanding concerning the genes associated with the development of the migratory phenotype would allow for preferential release of samples with the migratory alleles.

Further studies of this genetic predisposition for migration in rainbow trout has led to the discovery of an association between variation in migratory strategies and polymorphisms within the *GREB1-L* gene in both *Oncorhynchus mykiss* and *Oncorhynchus tshawytscha* from Oregon and California (Prince et al., 2016). Another study published in 2016 found 3 SNPs that were significantly associated with migration all residing within the *GREB1-L* gene (Hess et al., 2016). This variation within *GREB1-L* is good evidence for this gene being a key regulator of the migratory decision in *Oncorhynchus mykiss* (Prince et al., 2016). Although this association is known, the polymorphic positions leading to this variation between phenotypes is not. Knowing these polymorphic positions and the influence they have on the gene may help with the characterization of *GREB1-L* as a master control gene for migration within rainbow trout (Pearse et al., 2014). Moreover, the associations between *GREB1-L* and the decision to migrate have only been studied in two populations of rainbow trout. The importance of this gene in determining the migratory phenotype is clear in those two populations but remains to be found in other populations of rainbow trout.

Therefore, this project will look at a population of rainbow trout from Sashin Creek, Alaska for similar associations between *GREB1-L* and the decision to migrate. The Sashin Creek population has become a model system for answering questions concerning the genetic basis of migration in

rainbow trout, as QTL, GWAS, and gene expression studies all suggest migration in this population has a strong genetic component (Hecht et al., 2012; Hale et al., 2013; McKinney et al., 2015; Hale et al., 2016). Using this knowledge, the main aim of this project is to measure associations between polymorphisms in *GREB1-L* and the development of different migratory phenotypes in rainbow trout.

Methods

Sampling

All 1579 samples used in this study were taken from two crosses initiated in May of 2010 in Sashin Creek, Alaska: 1) between two mature steelheads that had returned to Sashin Creek to breed (M x M) and 2) between two mature residents that had remained in Sashin Creek (R x R). The samples were F1 from these crosses, or the first generation of offspring after the cross was initiated. After two years of letting this F1 population of fish grow, fin clip samples were taken from various individuals from each of the crosses and their phenotypes were identified. The migratory decision was categorized based upon the appearance of the trout with each sample either categorized as a mature resident male, an immature resident, an indeterminate, or a smolt (Figure 1).

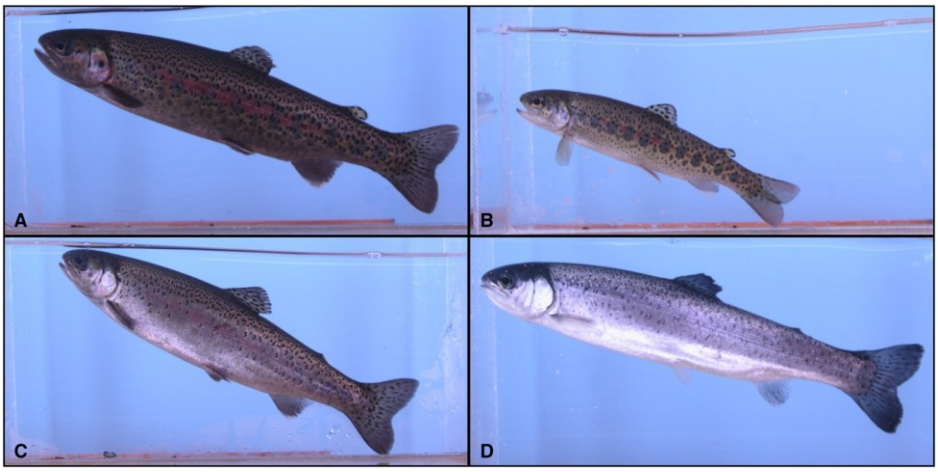

Figure 1: Examples of the four possible phenotypes used to categorize the F1 samples. A: Mature resident males. B: Immature resident. C: Indeterminate. D: Smolt (Hecht et al., 2015)

Sexing using OmyY1

DNA was extracted from the fin clips using a Qiagen extraction following manufacturers protocols. The extracted DNA was used to sex 156 M x M samples and 78 R x R samples. Three-hundred forty-nine samples from these crosses were not used because DNA was not extracted from their tissue or the samples came from an M x R cross that was not used in this study. Four-hundred one samples were mature males (i.e., produced gametes) so sexing them was not necessary, and 595 samples were not used as their phenotype was indeterminate (i.e., could not be classified as a resident or a smolt). Using designed primers, the *OmyY1* gene was amplified as it is found in the male specific region of the Y-chromosome (Brunelli et al., 2008). Amplification of *OmyY1* was performed in 10 µl reactions consisting of 2 µl of DNA, 5 µl of GoTaq, 2.8 µl of nuclease free water, and 0.1 µl of both the forward and reverse *OmyY1* primers. The PCR protocol was 94 °C denaturation temperature for 50 seconds, 55 °C annealing for 50 seconds, 72 °C elongation for 1 minute, and this cycle was repeated 35 times. The PCR product

was then run on an agarose gel and viewed using UV light to identify each sample as male or female (Figure 2).

To verify the results of this sexing procedure, a double-blind interpretation of the gels was conducted on approximately 50 samples from 5 different gels. Additionally, all the gels were run with a negative control to insure the PCR products were of reliable quality.

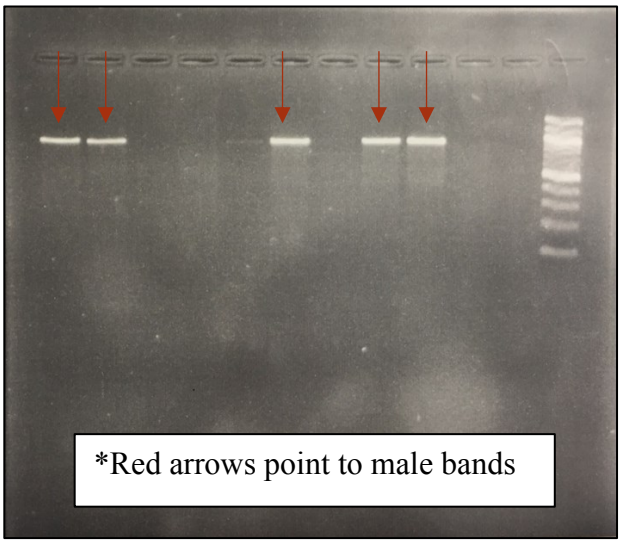

Figure 2: An agarose gel showing the PCR products from the *OmyY1* protocol. Lanes with a prominent, single band indicate a male and lanes without a band indicate a female.

Sequencing the *GREB1-L* Gene

Various samples were chosen for sequencing of the *GREB1-L* gene in groups of eight, including four smolts, two residents, and two mature males. Multiple primer pairs were made for the *GREB1-L* gene, but four of these primer pairs were specifically designed using a 2 million base pair scaffold of the *GREB1-L* gene constructed in another study (Prince et al., 2016). The A1 pair had to be discarded as the PCR product was too small for sequencing. Each of the remaining three primer pairs (Table 2) was around 600 base pairs in length, with each pair targeting a

different region of the scaffold. The study that constructed this *GREB1-L* scaffold found multiple SNPs within each of these regions within populations of rainbow trout from the Northern Umpqua tributary in Oregon and Eel River in Northern California (Table 1). These same regions were sequenced in the samples taken from Sashin Creek and later analyzed for polymorphisms between the different phenotypes.

The other primer pairs that were not designed from the scaffold were made using a *GREB1-L* reference sequence. The PCR products were not put through the sequencing process as many of the samples did not amplify and the ones that did amplify were not of good quality.

The PCR protocol for each of the three successful primers designed from the scaffold was 94 °C denaturation for 1 minute, varying annealing temperatures (A1.1=63 °C; A2=64 °C; A3=56 °C) for 1 minute, 72 °C elongation for 1 minute, and this cycle was repeated 40 times. The PCR products were then run on an agarose gel to insure their quality and purity (Figure 5). After confirming the PCR products were good enough for sequencing, an ExoSap procedure was performed in 7 µl reactions consisting of 5 µl of the product, 1.1 µl of nuclease free water, 0.2 µl of rSAP, 0.7 µl of NEB2 buffer, and 0.01 µl of Exo. The PCR protocol for this ExoSap procedure was 37 °C for 15 minutes and 80 °C for 15 minutes.

The Exo product was then used in a 10 µl Big Dye reaction consisting of 2 µl of the product, 5.45 µl of nuclease free water, 0.5 µl of Bright Dye (MCLAB), 1.75 µl of 5X sequencing buffer, and 0.3 µl of primer. Reactions were conducted for both the forward and reverse primers for each of the three primer pairs. The PCR protocol for the sequencing reaction was 96 °C

denaturation for 10 seconds, 50 °C annealing for 5 seconds, 60 °C elongation for 2 minutes and 30 seconds, and this cycle was repeated 30 times.

The sequencing products were then electrophoresed on an Applied Biosystems' 3130xl Genetic Analyzer. The sequenced samples were then analyzed using Sequencher version 5.1, allowing us to look for polymorphisms between the different samples.

Table 1: Spreadsheet of the 16 SNPs found within the scaffold from the Prince et al. study. The SNPs highlighted in blue represent the 13 SNPs targeted by the three primer pairs (Table 2) designed for the Sashin Creek population of rainbow trout used in this study.

GREB1-L Position	Migrant Allele	Resident Allele
568978	T	C
592595	G	A
592596	G	T
592627	T	C
595076	T	C
595084	A	G
595186	G	T
595253	A	G
649195	A	T
649286	T	C
649350	T	A
649428	A	G
649467	G	A
649544	A	C
780205	G	T
780229	T	C

Table 2: Primer pairs designed for the *GREB1-L* gene in the rainbow trout samples taken from Sashin Creek. The primers were designed using a 2 million base pair scaffold containing the *GREB1-L* gene.

Gene	Left Primer	Right Primer	Size	Location in Sequence
GREB1_A2	GTGGCCACTGCTTCAACTGT	TGATACAGTGAAATAATCTG	594	649062;649656 (scaffold)
GREB1_A3	AGCGTGAACACTCCAAAGGA	CTGAACTTGCTTCCCGACTC	650	594922;595551 (scaffold)
GREB1_A1.1	GGAGCTGACCCTGTTCTCTG	GACTGGGTCCCTCACACCTA	665	592277;592940 (scaffold)

Results

Sexing

Results from the sexing procedure show that there were more female migrants than male migrants in both the M x M cross and R x R cross (Figure 3). This difference between the sexes is similar to what was found in previous studies on rainbow trout due to inherited differences in life-history strategies between males and females (Hecht et al., 2012). The reason that we see this continuous difference between the sexes is that female fitness is more dependent on body size than male fitness because egg production increases with overall body size (Jonsson et al., 1998). In other words, the increased body size that migrants gain while living in the ocean (due to more resources) benefits the overall fitness of females more than males.

Sequencing

The polymorphic positions found in the Prince study were analyzed in a subset of our samples from Sashin Creek. Figure 4 shows PCR products that yielded the high-quality bands required to obtain good sequence data. Fifteen migrant and 15 resident F1 samples from the M x M cross were analyzed in all three regions of the scaffold. It should also be noted that only about 4.5% (1,800 bp surveyed / 40,000 total bp in the *GREB1-L* gene) of the *GREB1-L* gene was surveyed and analyzed for polymorphic positions (Prince et al., 2016; Hess et al., 2016). The previous study found 16 SNPs that had migrant associated alleles and resident associated alleles (Prince et al., 2016). Thirteen of these 16 positions were sequenced and analyzed in the Sashin Creek

population. No SNPs were found at any of these positions or any other position analyzed. However, at each of the 13 positions, all Sashin Creek samples, regardless of their migratory phenotype, showed the resident allele that was found in the previously studied populations (Prince et al., 2016; Table 3).

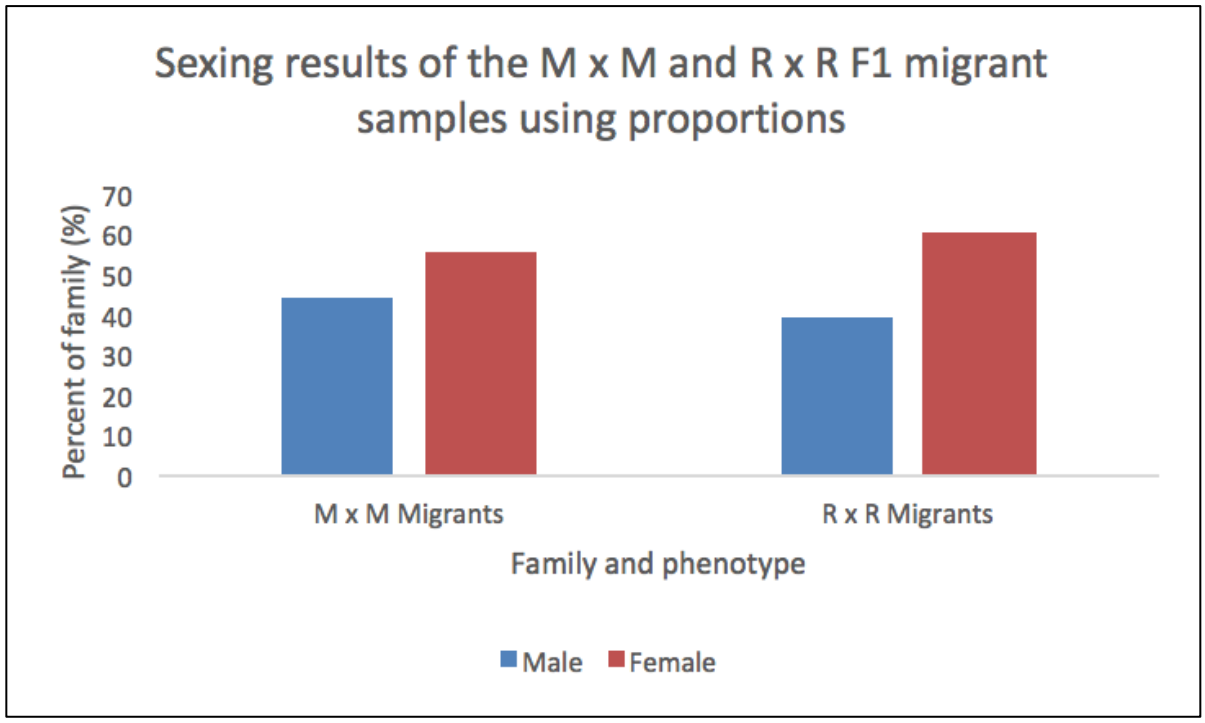

Figure 3: The percentage of male and female migrants in both the M x M cross and the R x R cross.

Table 3: The M x M F1 alleles from the Sashin Creek population at the positions described in Prince et al. (2016). Both migrants and residents from Sashin Creek showed the resident allele from Prince et al. (2016) at all 13 positions analyzed.

M x M F1 Samples			
GREB1-L Position	Sashin Creek Resident Allele	Sashin Creek Migrant Allele	Resident Allele (Prince et al., 2016)
592595	A	A	A
592596	T	T	T
592627	C	C	C
595076	C	C	C
595084	G	G	G
595186	T	T	T
595253	G	G	G
649195	T	T	T
649286	C	C	C
649350	A	A	A
649428	G	G	G
649467	A	A	A
649544	C	C	C

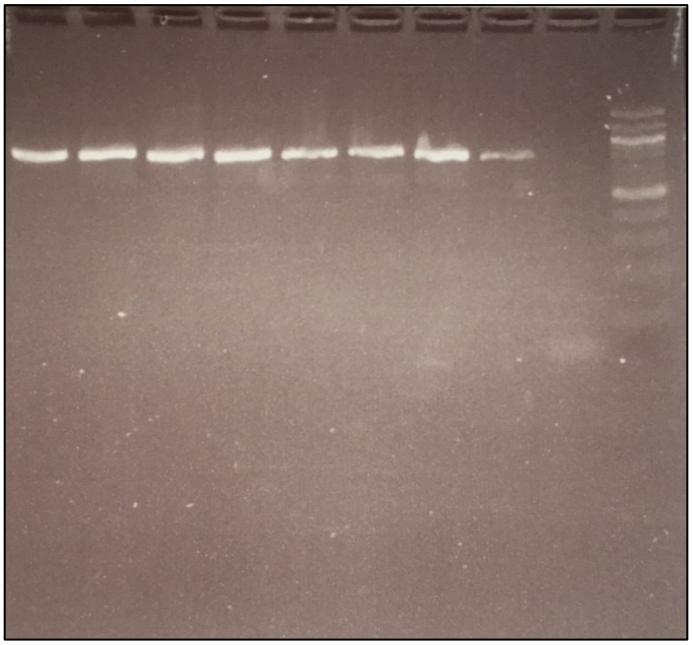

Figure 4: An agarose gel showing bands amplified using the A2 primer pair at 64 °C annealing temperature that gave good sequencing results.

Discussion

Previous studies have found a strong association between alleles within the *GREB1-L* gene and the propensity to migrate in rainbow trout. In this study, we did not find such an association with both migratory phenotypes being produced with the same alleles (i.e., no SNPs). This is good evidence for locally adapted genes like *GREB1-L* playing a key role in the migratory decision in *Oncorhynchus mykiss.* This local adaptation of *GREB1-L* within the rainbow trout genome is likely due to the different environmental factors affecting the different populations of rainbow trout. Ultimately, this all means that the different environments in which rainbow trout live may have different effects on the development of migration in *Oncorhynchus mykiss*. Additionally, these results also confirm previous findings that there are population-specific genetic effects that influence the migratory life history of rainbow trout. In support of this finding, previous studies

have shown the presence of loci that are not conserved between different populations of rainbow trout and that different genetic mechanisms determine the migratory life history in different populations of rainbow trout (e.g. Hale et al., 2013; Hecht et al., 2013). Comparing the population of rainbow trout used in other studies (i.e. Hecht et al., 2013) to the population of rainbow trout from Sashin Creek used in this study (i.e. Hale et al., 2013) confirms that different areas of the genome are associated with migration. Therefore, the lack of SNPs within *GREB1-L* may simply a reflect population-specific importance of this gene in the decision to migrate in rainbow trout.

All positions in *GREB1-L* that were polymorphic in other populations contained the resident allele in all Sashin Creek samples. This hints towards the probability that the development of the migratory phenotype in Sashin Creek may not have a strong association with polymorphisms in the *GREB1-L* gene as it did with other populations of rainbow trout (Prince et al.; 2016; Hess et al., 2016). However, only a small portion of the *GREB1-L* gene was sequenced and analyzed in this study (about 4.5% of the whole gene). Sequencing the remainder of the *GREB1-L* gene in the Sashin Creek population may yield some polymorphisms between the migratory phenotypes that can then be genotyped. If polymorphisms in other regions of the gene are found and genotyped, this may further strengthen the association between this gene and the decision to migrate in rainbow trout.

The findings from this project may also benefit from the sequencing and genotyping of different populations of rainbow trout for *GREB1-L*. Results from this project may indicate that populations at different latitudes have different allelic makeup at various locations along the

GREB1-L gene. Similar findings have been made between different latitudes of *Oncorhynchus mykiss* and an inversion in chromosome 5. Lower latitude populations contain the inversion whereas high latitude populations do not (Pearse et al., 2014). Perhaps, polymorphisms within the *GREB1-L* gene show similar variations with latitude. Incorporating a variety of populations from different latitudes would help to test this idea.

Understanding the genetic basis for migration in rainbow trout has been a topic of interest among researchers for some time now. It has also proven challenging. Finding out exactly how this migratory decision is inherited, the genetic predisposition for this decision, and how the environment impacts this decision are questions that still need to be answered. However, the findings from this project get us one step closer by allowing us to know more about the inheritance of migratory strategies in rainbow trout. This knowledge, in turn, may help us to form some conservation efforts for this species of migratory fish. In fact, conservation efforts may become a necessity in the future as we are beginning to see declines in *Oncorhynchus mykiss*, specifically steelhead trout, in the lower 48 states (Busby et al., 1996).

Additionally, the results found in this project may carry over to other species of migratory trout and salmon. A study published a year ago found a similar association between *GREB1-L* and the decision to migrate in Chinook salmon, proving that additional associations between genes and the migratory decision may be found in other migratory fish species (Prince et al., 2016). These associations found in this study and previous studies indicate that we are getting closer and closer to narrowing the gap between genetics and migration in all migratory fish species.

References

Bjornsson, Bjorn Thrandur, Sigurd O. Stefansson, and Stephen D. McCormick. "Environmental Endocrinology of Salmon Smoltification." *General and Comparative Endocrinology* 170, no. 2 (JAN 15, 2011): 290-298.

Brunelli, Joseph P., Kelsey J. Wertzler, Kyle Sundin, and Gary H. Thorgaard. "Y-Specific Sequences and Polymorphisms in Rainbow Trout and Chinook Salmon." *Genome* 51, no. 9 (SEP, 2008): 739-748.

Busby, P. J., T. C. Wainwright, G. J. Bryant, L. J. Lierheimer, R. S. Waples, F. W. Waknitz, I. V. Lagomarsino. "Status Review of West Coast Steelhead from Washington, Idaho, Oregon, and California." U.S. Dept. of Commerce, *NOAA* Tech. Memo., NMFS-NWFSC-27 (1996), 261 p.

Goetz, Fred A., Eric Jeanes, Megan E. Moore, and Thomas P. Quinn. "Comparative Migratory Behavior and Survival of Wild and Hatchery Steelhead (Oncorhynchus Mykiss) Smolts in Riverine, Estuarine, and Marine Habitats of Puget Sound, Washington." *Environmental Biology of Fishes* 98, no. 1 (JAN, 2015): 357-375.

Hale, Matthew C., Garrett J. McKinney, Frank P. Thrower, and Krista M. Nichols. "RNA-Seq Reveals Differential Gene Expression in the Brains of Juvenile Resident and Migratory Smolt Rainbow Trout (Oncorhynchus Mykiss)." *Comparative Biochemistry and Physiology.Part D, Genomics & Proteomics* 20, (2016 Sep 4 (Epub 2016 Sep 04), 2016): 136-150.

Hale, Matthew C., Frank P. Thrower, Ewann A. Berntson, Michael R. Miller, and Krista M. Nichols. "Evaluating Adaptive Divergence between Migratory and Nonmigratory Ecotypes of a Salmonid Fish, Oncorhynchus Mykiss." *G3-Genes Genomes Genetics* 3, no. 8 (AUG 1, 2013): 1273-1285.

Hecht, Benjamin C., Nathan R. Campbell, Dean E. Holecek, and Shawn R. Narum. "Genome-Wide Association Reveals Genetic Basis for the Propensity to Migrate in Wild Populations of Rainbow and Steelhead Trout." *Molecular Ecology* 22, no. 11 (JUN, 2013): 3061-3076.

Hecht, Benjamin C., Jeffrey J. Hard, Frank P. Thrower, and Krista M. Nichols. "Quantitative Genetics of Migration-Related Traits in Rainbow and Steelhead Trout." *G3-Genes Genomes Genetics* 5, no. 5 (MAY 1, 2015): 873-889.

Hecht, Benjamin C., Frank P. Thrower, Matthew C. Hale, Michael R. Miller, and Krista M. Nichols. "Genetic Architecture of Migration-Related Traits in Rainbow and Steelhead Trout, Oncorhynchus Mykiss." *G3-Genes Genomes Genetics* 2, no. 9 (SEP 1, 2012): 1113-1127.

Hess, Jon E., Joseph S. Zendt, Amanda R. Matala, and Shawn R. Narum. "Genetic Basis of Adult Migration Timing in Anadromous Steelhead Discovered through Multivariate

Association Testing." *Proceedings of the Royal Society B-Biological Sciences* 283, no. 1830 (MAY 11, 2016): 20153064.

Jonsson, N., B. Jonsson, and LP Hansen. "Long-Term Study of the Ecology of Wild Atlantic Salmon Smolts in a Small Norwegian River." *Journal of Fish Biology* 52, no. 3 (MAR, 1998): 638-650.

Laura Liberoff, Ana, Analia Pamela Quiroga, Carla Marcela Riva-Rossi, Jessica Adele Miller, and Miguel Alberto Pascual. "Influence of Maternal Habitat Choice, Environment and Spatial Distribution of Juveniles on their Propensity for Anadromy in a Partially Anadromous Population of Rainbow Trout (Oncorhynchus Mykiss)." *Ecology of Freshwater Fish* 24, no. 3 (JUL, 2015): 424-434.

McKinney, Garrett J., Matthew C. Hale, Giles Goetz, Michael Gribskov, Frank P. Thrower, and Krista M. Nichols. "Ontogenetic Changes in Embryonic and Brain Gene Expression in Progeny Produced from Migratory and Resident Oncorhynchus Mykiss." *Molecular Ecology* 24, no. 8 (APR, 2015): 1792-1809.

Nichols, Krista M., Alicia Felip Edo, Paul A. Wheeler, and Gary H. Thorgaard. "The Genetic Basis of Smoltification-Related Traits in Oncorhynchus Mykiss." *Genetics* 179, no. 3 (JUL, 2008): 1559-1575.

Pearse, Devon E., Michael R. Miller, Alicia Abadia-Cardoso, and John Carlos Garza. "Rapid Parallel Evolution of Standing Variation in a Single, Complex, Genomic Region is Associated with Life History in steelhead/rainbow Trout." *Proceedings of the Royal Society B-Biological Sciences* 281, no. 1783 (MAY 22, 2014): 20140012.

Prince, Daniel J., Sean M. O'Rourke, Tasha Q. Thompson, Omar A. Ali, Martha Arciniega, Hanna S. Lyman, Ismail K. Saglam, Anthony J. Clemento, Thomas J. Hotaling, Andrew P. Kinziger, Adrian P. Spidle, John Carlos Garza, Devon E. Pearse, Michael R. Miller. In Review. "The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation." bioRxiv doi: 10.1101/056853.

Thrower, FP, JJ Hard, and JE Joyce. "Genetic Architecture of Growth and Early Life-History Transitions in Anadromous and Derived Freshwater Populations of Steelhead." *Journal of Fish Biology* 65, (DEC, 2004): 286-307.