INFLUENCE OF LANDSCAPE AND ENVIRONMENT ON SALMONID GENETICS

Shawn R. Narum Nathan R. Campbell Andrew P. Matala Ben Hecht Jon Hess, Columbia River Inter-Tribal Fish Commission (CRITFC), Hagerman, ID, 83332

Table of Contents

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

This project addresses two objectives related to environmental and landscape features that contribute to population structure, life history diversification, and adaptation of salmonids.

Objective 1) Environment & Landscape Genetics – Evaluate genetic structure of natural populations of salmonids relative to their environment and identify candidate markers associated with traits that are related to adaptation of steelhead and Chinook salmon populations

During the performance period of July 1, 2013 to June 30, 2014, work was completed that addresses components of both Objective 1 and Objective 2. For Objective 1, an extensive genetic baseline with 192 SNPs was created for steelhead to evaluate neutral and adaptive genetic variation of populations throughout the Columbia River Basin (Section 2). Work has also progressed on sequencing Chinook salmon throughout the Columbia River Basin to evaluate adaptive genetic variation related to environmental features (Section 3).

Objective 2) Controlled Experiments – experiments with controlled environmental variables to validate phenotypic response of fish with given genotypes.

For Objective 2, empirical work was done to further advance our understanding multiple traits related to recovery of salmonids in the Columbia River. Investigations during the past performance period focused on patterns of gene expression of various strains of redband trout under heat stress (Section 1), run-timing in Chinook salmon and steelhead (Section 3). Further, progress was also made towards developing projects to investigate the genomic basis for age-at-maturity in Chinook salmon and thermal tolerance in *O. mykiss* (Section 3).

Acknowledgements

Report Structure

This report is divided into three sections. The first section reports on landscape genetics and adaptive variation in Columbia R. steelhead, the second section on patterns of gene expression of various strains of redband trout under heat stress, and the third section provides brief summaries for ongoing and future work on several traits of interest including heritability and genetic basis of age-at-maturity in Chinook salmon, genetic basis of run-timing in Chinook salmon and steelhead, and thermal adaptation in O. mykiss.

SECTION 1: Transcriptomic response to heat stress among ecologically divergent populations of redband trout

Introduction

Thermal adaptation is a widespread phenomenon in organisms that are exposed to variable and extreme environments. While some organisms may alter their distribution or behavior to avoid stressors and others may acclimate through physiological plasticity [1,2], many species evolve adaptive responses to local conditions over generations through natural selection [3-5]. Evolutionary adaptation to local environments has been demonstrated across a wide variety of taxa [6], and is expected to play a critical role for species with limited dispersal capabilities. However, few studies have identified the underlying molecular mechanisms that have led to conspecific adaptation to thermal conditions.

Molecular techniques such as RNA-seq [7] provide the opportunity to investigate transcriptional response to thermal stress and further identify mechanisms for thermal adaptation. Patterns of gene expression under heat stress are important to determining evolutionary adaptation among conspecific populations that occupy various environments. Multiple genes have been shown to be involved in heat tolerance across many species, including highly conserved heat shock proteins (hsps) that are upregulated under stressful conditions such as exposure to heat [8,9]. An adaptive heat shock response has additionally been shown to occur among conspecific populations that occupy variable environments [3,10]. However, many genes are known to have a role in regulating the effects of temperature and are likely to be involved in thermal adaptation [11,12]. Thus, RNA-seq provides the opportunity to investigate differential expression across the transcriptome and identify biological pathways involved in evolutionary response to thermal stress.

Redband trout (*Oncorhynchus mykiss gairdneri*) occupy highly variable environments including both montane and desert streams and have been shown to be locally adapted to these different environments [13]. Previous research has demonstrated an adaptive heat shock response among populations from different climates but also suggests that additional mechanisms are involved with thermal adaptation [14]. This species appears to have evolved a finely tuned heat shock response that likely requires additional genes to balance the short term (immediate cellular damage) and long term (fitness) costs associated with thermal stress. Given that oxygen delivery is limiting for fish under climate-related stressors [15], genes involved in oxygen transport are expected to play a significant role. Additionally, we expect that metabolic and immune pathways could be involved given the energy demands and potential for disease under chronic environmental stress [16,17]. Therefore, this study tests for molecular response to heat stress across the transcriptome of ecologically divergent populations of redband trout that have evolved under local climate regimes.

In this study, desert and montane strains of redband trout and their F1 crosses were tested for differential gene expression under heat stress in a common garden experiment. We tested for both acute and chronic stress response by quantifying gene expression in fish that experienced diel water temperatures similar to desert environments that peaked near their thermal maxima (~28.5 C) over the course of four weeks. Tissues were collected from each strain of redband trout at multiple time points to test for both acclimatization within each strain and evolutionary adaptation between strains. Results are quantified to confirm the role of hsp genes in these fish, but also identify additional genes and biological pathways that are regulated to balance the costs of stress response in populations that have evolved to desert environments.

Methods

Redband trout populations and thermal stress experiments

To investigate thermal acclimation and adaptation of redband trout (*O. mykiss gairdneri*) from desert and montane populations, fry of approximately four months of age from each environment and their F1 crosses were exposed to diel temperature cycles (peaking at 28°C) over a 4-week period in a controlled setting. Gill tissues were collected from euthanized individuals on day 1, 3, 7 and 28 to quantify mRNA expression across the transcriptome.

Gametes and fry were collected from two ecologically divergent populations: one from a desert climate stream Little Jacks Cr. (LJ), and one from a montane climate stream Keithley Cr. (K), both located in Idaho, USA. In order to create F1 crosses, gametes from each strain were cross fertilized and reared in laboratory incubators. The two sites were chosen for study based on previous tests of redband trout from six desert and six montane streams that demonstrated that Little Jacks Cr. fish were adapted to a desert climate and Keithley Cr. was a typical montane stream population (Narum et al. 2010). Gametes were fertilized to produce half-sibling progeny representing three distinct strains: one of pure desert strain (LJ), one of pure montane strain (K), and the F1 crosses (LJ males x K females). Fry were reared in constant 15 °C spring water until they reached an average weight of 2 g, then each strain was divided into treatment and control groups. Three replicate tanks were used for all treatment and control groups for each strain (3 strains x 2 treatments x 3 replicates equals a total of 18 tanks) with an average of 45 fish per tank. Fish were fed a diet of Soft Moist pellets (Rangen Inc.) to satiation twice per day, and photoperiod was fixed at 14 h light and 10 h darkness. Fish in recirculating treatment tanks experienced diel temperature cycles over 6 weeks that reached a maximum of 28.5 °C in the afternoon and a minimum of 17.0 °C at night (mean temperature gradient of ~1.5 °C per hour;

Fig. S1, Supporting information), while fish in control tanks were held at a constant temperature of 15 °C spring water. All experimental protocols were approved in advance by the University of Idaho's Institutional Animal Care and Use Committee (Protocol #201025).

RNA-seq library prep and Illumina sequencing

Total RNA was isolated from approximately 5 mg of gill tissue from individual fish using Qiagen RNeasy kits. RNA was normalized to 100 ng/µL and equal volumes of RNA from three fish from each tank were pooled for a total of 72 libraries (18 tanks x 4 time periods each; Table 1). The Ribo-Zero™ Magnetic Gold Kit (Epicentre) was used to deplete the samples of ribosomal RNA (rRNA) which constitutes a large proportion of the total RNA. Ribosomal RNA depleted samples were purified using the ethanol precipitation method suggested in the Ribo-Zero Kit protocol and resuspended in 20µL of RNAse free water. RNA-seq libraries were prepared using 4.75µL of template material using the strand-specific library preparation kit ScriptSeq™ v2 RNA-Seq Kit (Epicentre). The tagged cDNA constructs were purified using Qiagen MinElute columns and sample specific index sequences were added during PCR [95ºC – 1m; (95ºC – 30s; 55ºC – 30s; 68ºC – 3m) x 15; 68ºC – 7m; 4ºC – hold] using ScriptSeq (Epicentre) index PCR primers. Each amplified library was then purified using Agencourt AMPure XP magnetic beads and eluted with 20uL of nuclease free water. Dilutions (1:2000) of the purified and indexed libraries were then quantified by qPCR using a ABI-7900 instrument (Life Technologies), Power-Sybr master mix (Life Technologies), standard Illumina (P5,P7) primers and an Illumina PhiX library as a standard. The indexed libraries were normalized to 10nM concentration in Tris-EDTA (pH 8.0) buffer with 0.1% Tween-20 and combined for sequencing (8 pooled libraries with nine samples, each with ScriptSeq index sequences 1-9).

Each of the pooled libraries was sequenced in two lanes of a single read 100bp flow cell on an Illumina HiSeq 1500 instrument for a total of 16 lanes. Each lane of data was demultiplexed by index sequence and reads were combined from both lanes for each sample. The average number of reads per sample after quality filtering was 31.96 M and ranged from 22.40 – 72.70 M. Raw data was submitted to NCBI's short read archive (SRA; entry GSE53907).

Sequence alignment to reference transcriptome

Raw sequencing data was aligned to a reference transcriptome for rainbow trout designed for stress response [18] using the program Bowtie [19]. Parameters for Bowtie were set to exclude the first 10 bases and last 30 bases of sequencing data leaving 60 bases of high quality sequence for alignment. Both the forward (+) and reverse complement (-) of each reference transcript were considered for alignment since the reference transcriptome was assembled from non-directional sequence data. The best single match to the reference transcriptome was returned provided it had no more than 3 mismatches across the 60 bases (95% identical).

Output data from Bowtie was condensed by counting matches to each of the reference transcript contigs for both + and - orientation. Since the library preparation process generates directional constructs (strand-specific library preparation), we expected legitimate alignments to match only one of the two orientations of each reference transcript. Indeed, we found that our library reads predominantly matched only 1 of the 2 strand orientations. Therefore, read alignments to the minor strand were considered mis-assigned and excluded from the data set. Finally, since the reference transcriptome included several sequence variants (contigs) for many putative genes, combined read counts from contigs within each gene were utilized for differential expression analyses.

Differential expression analyses

Tests for differential expression were completed using the edgeR Bioconductor package [20]. Differential expression analyses were done to test for two processes involved in heat stress response: 1) acclimatization to chronic heat stress over time, and 2) evolutionary adaptation of the specific strains. To test for acclimatization, differential expression was tested separately for each time period (Day 1, 3, 7, 28) with all strains in the model. To test for evolutionary adaptation, differential expression was tested separately for each strain (LJ, K, LJxK) with all time periods in the model. Each model included the additional factor of condition (treatment or control) and three biological replicates. Genes that were not expressed in either condition were removed; specifically only genes with at least two counts per million reads in at least nine samples were kept for further analyses. Gene counts were normalized with the trimmed mean of 'M' values (TMM) method in edgeR as this has been shown to be one of the most reliable methods for this purpose in RNA-seq studies [21]. Sample metadata and normalized gene expression data was submitted to NCBI's Gene Expression Omnibus (GEO; entry GSE53907). As suggested by McCarthy et al. [22], genewise dispersion and a general linear model (GLM) were used for tests of differential expression. Genewise dispersion estimates deprioritize genes with inconsistent results and allow the main analysis to focus on changes that are consistent between biological replicates. The GLM accounts for the multifactor design of this study. A false discovery rate (FDR at 0.05; [23]) was applied to account for multiple tests of differentially expressed genes. In order to validate patterns of gene expression in the RNA-seq data, results were compared to quantitative PCR (qPCR) data for multiple heat shock genes. Significantly

regulated contigs in each strain were annotated in Bast2GO with a blastx minimum e-value set to 1.0E-06 [24].

Results

Sequence alignment to reference transcriptome

A total of 16 lanes for eight pooled libraries provided 2.30 billion quality filtered reads over all 72 samples. Read numbers ranged from 22.40-72.40 M per sample, with an overall mean of 31.96 M (Table 1). Read numbers were well balanced between treatment and control groups with 1.10 billion and 1.21 billion reads, respectively. Trimmed 60bp reads were aligned with a minimum criterion of 95% identity to the reference transcriptome at an average of 24.1% and 7.53 M reads per sample. Mean percent alignment and mean number of aligned reads for each set of biological replicates ranged from 11.1-40.9% and 3.32-12.49 M, respectively (Table S1). Reads aligned to a total of 25,128 unique contigs from the reference transcriptome. Principal component analyses of overall gene expression data clearly showed that samples clustered by treatment or control condition and also that distinct clusters were present for the desert and montane strains, but the cluster for the F1 strain overlapped with the montane strain (Fig. 1).

Figure 1: **Principal components analysis of overall transcriptome expression.** Results for 18 samples of redband trout collected either at first exposure to heat stress up to 28°C (darker shades) or on the same day from 15°C control temperature (lighter shades). Samples are color coded by their environment: desert strain (red; LJ=Little Jacks Cr.), F1 crosses (gray; LJxK), montane strain (blue; K=Keithley Cr.).
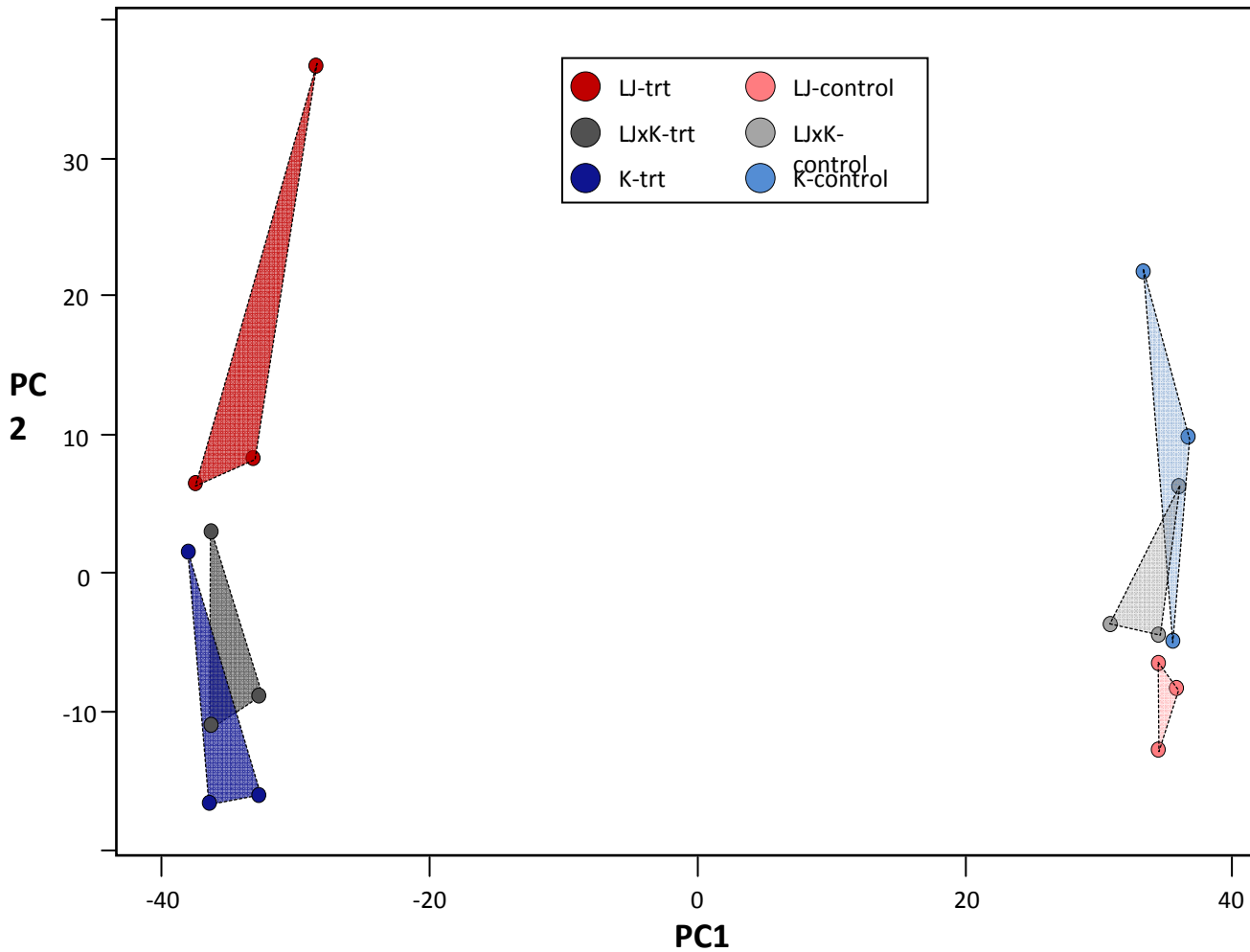
Differential expression

Differences in gene expression were highly significant within each strain (Figure 2a), with 7,051 significant genes for the desert strain (4,238 upregulated, 2,813 downregulated), 6,906 for F1 crosses (3,375 upregulated, 3,531 downregulated), and 6,774 for montane (3,499 upregulated, 3,275 downregulated). As expected, a large number of total transcripts (12,814) were differentially expressed in the study (FDR ≤ 0.05) with 2,310 transcripts occurring in common among all three strains, but the desert strain had a larger number of unique differentially expressed transcripts (2,875) than the montane (1,982) or the F1 (2,355) strain (Figure 2a). Strongly differentiated genes (> 4 fold change and FDR ≤ 0.05; Figure 2b) were particularly abundant in the desert strain (824 unique transcripts; Table S1) relative to the other two strains (montane = 58 Table S2; F1 = 192; Table S3). The desert strain also had a particularly large number of transcripts that were upregulated relative to the other two strains (Figures 3a-c).

Differences in gene expression were also observed over time as the number of significant genes consistently decreased with more days of exposure to temperature stress (Figure 4). The number of significant genes was 7,833 at Day 1 (4,058 upregulated, 3,775 downregulated), 6,408 at Day 3 (3,344 upregulated, 3,064 downregulated), 3,624 at Day 7 (1,958 upregulated, 1,666 downregulated), and 1,269 at Day 28 (719 upregulated, 550 downregulated). This trend was consistent with the expectation that the stress response would become reduced with chronic exposure to heat stress.

Gene ontology with Blast2GO revealed that strongly differentiated genes in each strain (> 4 fold change and FDR ≤ 0.05) included several categories for each of biological processes, molecular function, and cellular components (Figures 5a-c). Within biological process, there

were a total of 18 pathway categories at level 2 gene ontology, but nearly 70% of the genes were

included in five categories: cellular process (mean = 17.2%), metabolic process (mean = 16.7%),

Figure 2: **Venn diagrams of differentially expressed genes.** Results of each strain of redband trout for a) all significant transcripts (FDR ≤ 0.05); and b) strongly differentiated transcripts (> 4 fold change and FDR ≤ 0.05). Circles are color coded to represent fish by their environment: desert strain (red; LJ=Little Jacks Cr.), F1 crosses (gray; LJxK), montane strain (blue; K=Keithley Cr.).
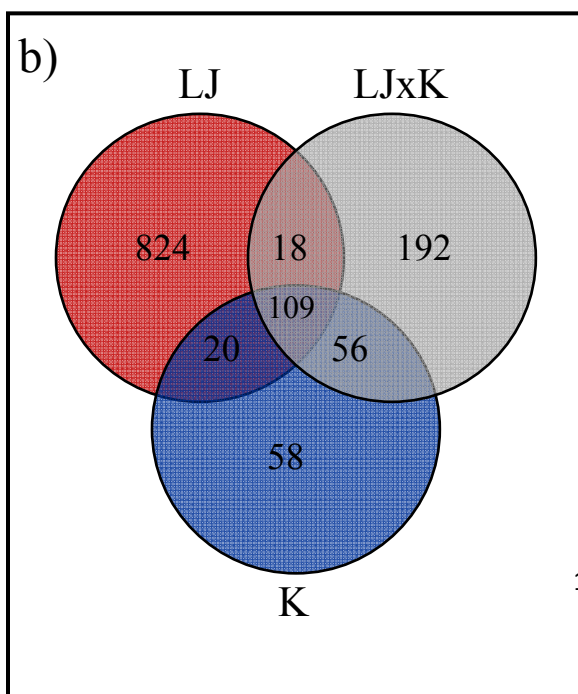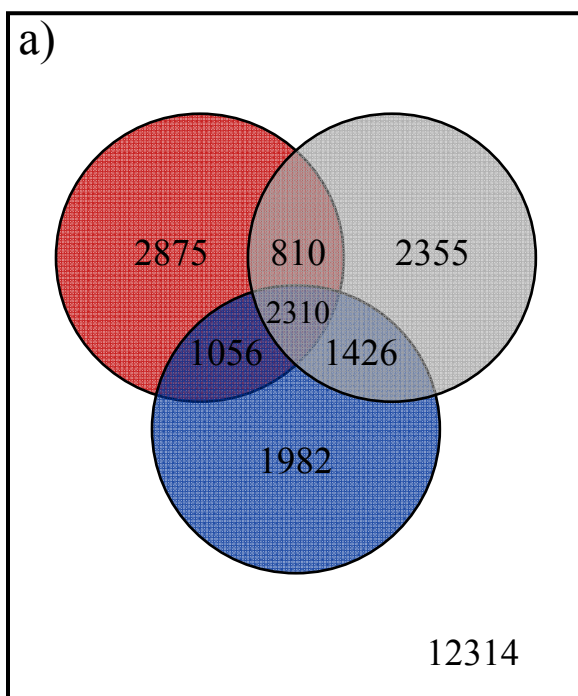
Figure 3: **Adaptation: differential expression for each strain of redband trout across all time periods of heat stress versus fish held at control temperatures.** Results for a) desert strain from Little Jacks Cr. b) F1 crosses; and c) montane strain from Keithley Cr. Genes that are significantly differentiated (FDR ≤ 0.05) are in red and those that are not significant are in black. On a log_2 scale, the green lines show genes that are ≥ 2 fold, and the blue lines designate genes that are ≥ 4 fold. The x-axis is the mean expression of each gene in counts per million reads (CPM) on a log_2 scale.
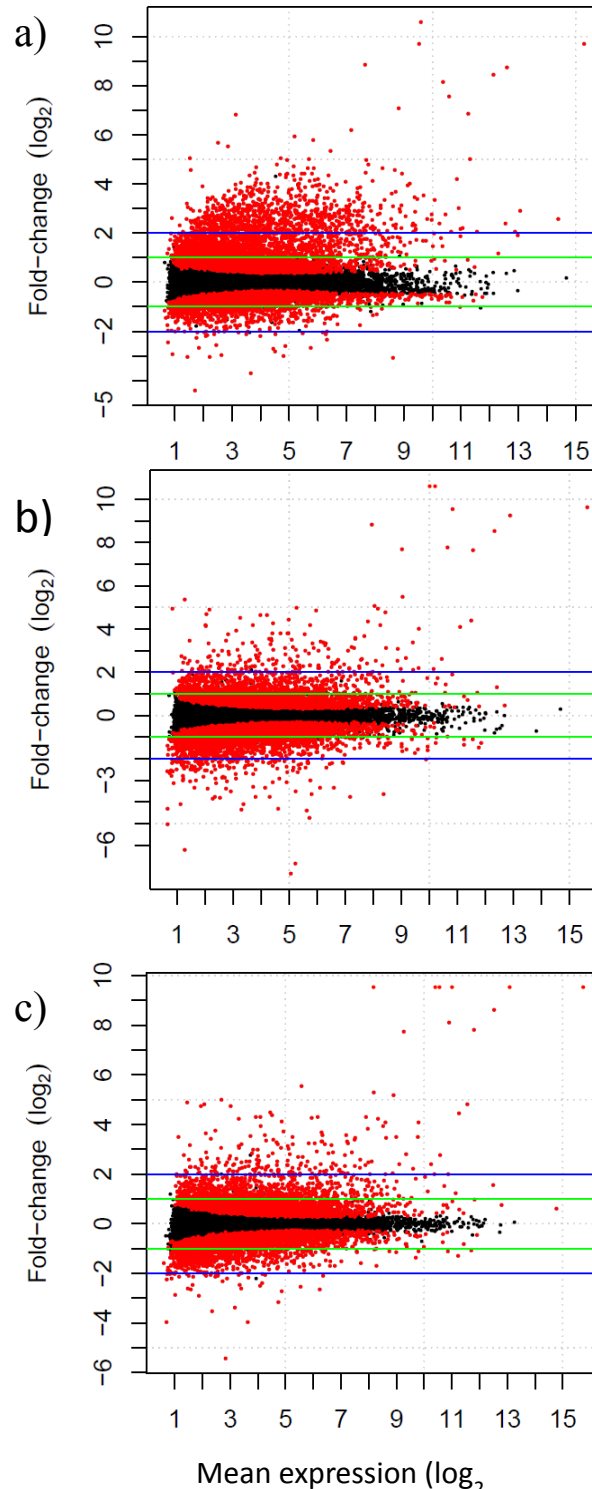
response to stimulus (mean = 12.2%), single-organism process (mean = 12.4%), and biological

regulation (mean = 10.3%; Figure 5a). Within molecular function, there were a total of 11

pathway categories at level 2 gene ontology with over 75% of the genes in two categories:

binding (mean = 52.3%), and catalytic activity (mean = 24.3%; Figure 5b). Within cellular

process, there were a total of 10 pathway categories at level 2 gene ontology, with over 80% of

the genes included in four categories: cell (mean = 31.3%), organelle (mean = 24.8%),

membrane (mean = 16.1%), and macromolecular complex (mean = 10.5%; Figure 5c).

Patterns of gene expression for each strain over time were compared with results from

qPCR assays for six heat shock genes and were highly consistent with either RNA-seq or qPCR

data. Specifically, expression patterns showed that heat shock genes were significantly lower for

the desert strain at Day 1 for all hsp genes and all strains had decreased gene expression from

Day 3 through the remainder of the experiment as shown previously (Narum et al. 2013).

Discussion

Since thermal stress has broad biological effects on organisms, transcriptional response is

expected to be highly diverse across several genes in ectothermic species such as redband trout.

This study confirms that numerous genes are differentially expressed in redband trout under heat

stress, and several pathways are involved. However, there were key pathways that contained a

large proportion of differentially expressed transcripts including response to stimulus, metabolic

processes, cellular processes, molecular binding function, and cell membrane function. These

pathways correspond well with previous studies that demonstrate these as critical physiological

components involved with of aquatic ectotherms exposed to elevated water temperatures [17,25].

In particular, several physiological studies have linked thermal tolerance with aerobic scope and

Figure 4: **Acclimation: differential expression at each time period of heat stress versus fish held at control temperatures across all strains.** Results for a) Day 1 of heat stress treatment; b) Day 3 of heat stress treatment; c) Day 7 of heat stress treatment; and d) Day 28 of heat stress treatment. Genes that are significantly differentiated (FDR $\leq$ 0.05) are in red and those that are not significant are in black. On a $\log_2$ scale, the green lines show genes that are $\geq$ 2 fold, and the blue lines designate genes that are $\geq$ 4 fold. The x-axis is the mean expression of each gene in counts per million reads (CPM) on a $\log_2$ scale.
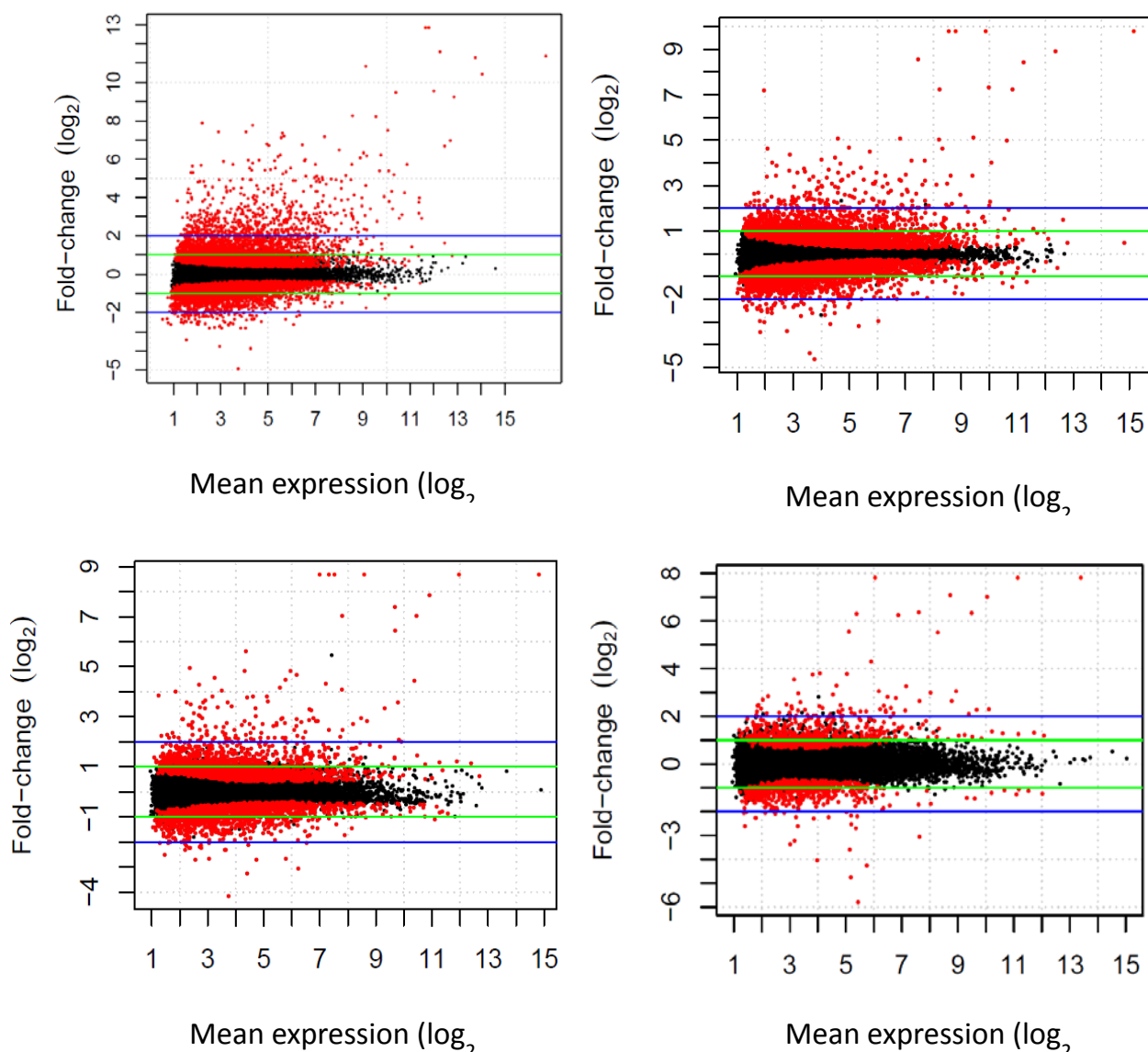
emphasize the role of metabolic processes in thermal adaptation (e.g., [25-27]). The larger number of differentially expressed genes in the desert strain versus the other two strains suggests that a complex combination of genes has evolved for redband trout in their desert environment.

Evidence for acclimation to heat stress was extensive as the number of differentially expressed transcripts decreased by 83.7% from Day 1 to Day 28. Results from this study elaborate on previous findings in redband trout that stress response genes are highly upregulated when exposed to heat stress [28,14]. Multiple heat shock genes (e.g., hsp70, hsp90, hsp47) were differentially expressed in all strains and time periods. However, an acclimation effect was evident as expression levels decreased over time in all strains. This is consistent with theories of acclimation to heat stress where organisms are able to moderate their heat shock response over time, as opposed to initial exposure where immediate survival is a priority [29].

More importantly, this study demonstrates that adaptive patterns of expression have evolved in ecologically divergent populations of this species. Results from Narum et al. [14] specifically highlight the adaptive response of heat shock genes in redband trout, with lower hsp gene expression observed in desert versus montane strains. Results in heat shock genes from the current RNA-seq data corroborate the previous qPCR results and emphasize that warm adapted natural populations are likely to have evolved a specialized heat shock response that reduces physiological costs of hsp production. This result is consistent with the adaptive heat shock response observed in natural populations of other organisms such as killifish (*Fundulus heteroclitus*; [30]) and *Drosophila buzzatii* [10]. This remains an important finding of this study and provides clarification regarding evolutionary adaptation of hsp gene expression in heat tolerant populations. However, many recent studies indicate that complex mechanisms are

involved in thermal adaptation of aquatic ectotherms beyond heat shock response (e.g., [25,31,32]. Indeed, this study of the transcriptome revealed adaptive patterns in metabolic and cellular process genes that suggest desert fish are more efficient at supporting these pathways than montane fish under heat stress, and chronic exposure may cause failure of these genes to be expressed in montane and F1 crosses and suggests that some critical physiological functions become limited in these strains over time. Previous studies suggest that metabolic pathways may be particularly important since metabolic energy stores are positively correlated with physiological function and swimming behavior in thermally adapted redband trout [16,17].

A variety of immediate and long-term anthropogenic disturbances such as habitat disturbance and climate change have negative impacts on freshwater fish [33,34] and the need to understand mechanisms for thermal adaptation in these organisms is critical. Many fishes have already been extirpated from large portions of their historical range (e.g., [35]) and the effects of climate change are expected to further alter species' range, phenology, and persistence [36-38]. Genomic and physiological mechanisms for thermal adaptation can be important tools for conservation measures to enable long-term viability of wild populations [39-41]. Specifically, this study helps to further identify tools such as genetic screening with candidate markers and field measurements of cardiac function that may be utilized to screen broadly across the species' range to predict the potential for adaptation under scenarios of climate change [42].

Conclusions

This study demonstrates that redband trout from a desert climate have a much larger number of strongly differentially expressed genes than montane and F1 strains in response to heat stress, suggesting that a complex combination of genes has evolved for redband trout in their desert

environment. Recent studies of physiological adaptation in aquatic ectotherms indicate that intraspecific thermal tolerance is set by limitations in aerobic performance, specifically the upper limit of heart rate to deliver more oxygen to tissues (e.g., [15,25,26]). This is due to temperature dependent oxygen limitation in aquatic environments, a theory that has been well supported in many organisms [43]. In order to support this increase in cardiac performance, redband trout would need to upregulate genes from multiple pathways including those observed in this study (e.g., metabolic pathways). However, further studies that specifically link allele specific gene expression with physiological functions such as aerobic scope and heart rate are needed to further elucidate the specific mechanisms involved with thermal adaptation in this species.

Figure 5: **Gene ontology (GO) annotation for transcripts that were strongly differentiated expressed in each strain (> 4 fold change and FDR ≤ 0.05).** Results shown for level 2 categories for a) Biological process; b) Molecular function; c) Cellular component. Bars are color coded to represent fish by their environment: desert strain (red; LJ=Little Jacks Cr.), F1 crosses (gray; LJxK), montane strain (blue; K=Keithley Cr.).
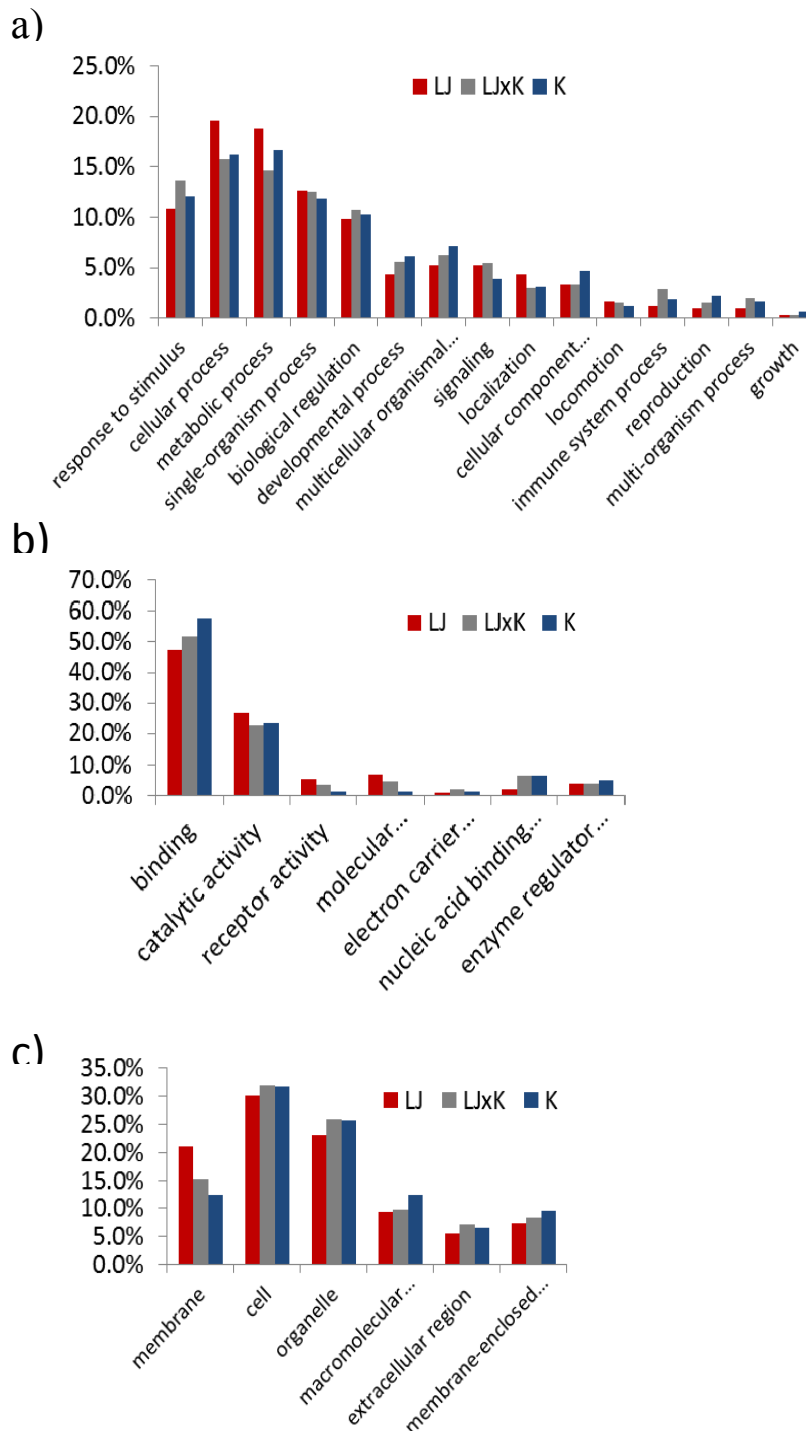
References

1. Angilleta MJ: *Thermal Adaptation: a Theoretical and Empirical Synthesis*. Oxford: Oxford University Press; 2009.

2. Whitehead A, Galvez F, Zhang S, Williams LM, Oleksiak MF: Functional genomics of physiological plasticity and local adaptation in killifish. *J Hered* 2011, 102: 499-511.

3. Dahlhoff EP, Rank NE: Functional and physiological consequences of genetic variation at phosphoglucose isomerase: Heat shock protein expression is related to enzyme genotype in a montane beetle. *Proc Natl Acad Sci* 2000, 97:10056-10061

4. Hoffman AA, Sorensen JG, Loeschcke V: Adaptation of Drosophila to temperature extremes: bringing together quantitative and molecular approaches. *J Therm Biol* 2003, 28:175-216.

5. Kavanagh KD, Haugen TO, Gregersen F, Jernvall J, Vollestad LA: Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. *BMC Evolutionary Biology* 2010, 10:350.

6. Keller I, Seehausen O: Thermal adaptation and ecological speciation. *Mol Ecol* 2012, 21:782–799.

7. Wolf JBW: Principles of transcriptome analysis and gene expression quantification: an RNA-seq tutorial. *Mol Ecol Res* 2013, 13:559–72.

8. Morimoto RI, Sarge KD, Abravaya K: Transcriptional regulation of heat shock genes. *J Biol Chem* 1992, 267:21987–21990.

9. Sørensen JG, Kristensen TN, Loeschcke V: The evolutionary and ecological role of heat shock proteins. *Ecol Lett* 2003, 6:1025–1037.

10. Sørensen JG, Dahlgaard J, Loeschcke V: Genetic variation in thermal tolerance among natural populations of *Drosophila buzzatii*: down regulation of Hsp70 expression and variation in heat stress resistance traits. *Funct Ecol* 2001, 15:289–296.

11. Sørensen JG, Norry FM, Scannapieco AC, Loeschcke V: Altitudinal variation for stress resistance traits and thermal adaptation in adult *Drosophila buzzatii* from the New World. *J Evol Biol* 2005, 18:829–837.

12. Kassahn KS, Crozier RH, Ward AC, Stone G, Caley MJ: From transcriptome to biological function: environmental stress in an ectothermic vertebrate, the coral reef fish *Pomacentrus moluccensis*. *BMC Genomics* 2007, 8:358.

13. Narum SR, Campbell NR, Kozfkay CC, Meyer KA: Adaptation of redband trout in desert and montane environments. *Mol Ecol* 2010, 19:4622-4637.

14. Narum SR, Campbell NR, Meyer KA, Miller MR, Hardy RW: Thermal adaptation and acclimation of ectotherms from differing aquatic climates. *Mol Ecol* 2013, 22:3090-9097.

15. Pörtner HO, Knust R: Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science* 2007, 315:95-97.

16. Gamperl AK, Rodnick KJ, Faust HA, Venn EC, Bennett MT, Crawshaw LI, Keeley ER, Powell MS, Li HW: Metabolism, swimming performance, and tissue biochemistry of

high desert redband trout (*Oncorhynchus mykiss* ssp.): evidence for phenotypic differences in physiological function. *Physiol Biochem Zool* 2002, 75:413–431.

17. Rodnick KJ, Gamperl AK, Lizars KR, Bennett MT, Rausch RN, Keeley ER: Thermal tolerance and metabolic physiology among redband trout populations in south-eastern Oregon. *J Fish Biol* 2004, 64:310–335.

18. Sánchez CC, Weber GM, Gao G, Cleveland BM, Yao J, Rexroad CR: Generation of a reference transcriptome for evaluating rainbow trout responses to various stressors. *BMC Genomics* 2011, 12:626.

19. Langmead B, Trapnell C, Pop M, Salzberg SL: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009, 10:R25.

20. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26:139-140.

21. Dillies M-A, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, et al: A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 2012, 10:1093.

22. McCarthy, DJ, Chen, Y, Smyth, GK: Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 2012, 40:4288- 4297.

23. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995, 57:289-300.

24. Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, Robles M: Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 2005, 21:3674–3676.

25. Eliason EJ, Clark TD, Hague MJ, Hanson LM, Gallagher ZS, Jeffries KM, Gale MK, Patterson DA, Hinch SC, Farrell AP: Differences in thermal tolerance among sockeye salmon populations. *Science* 2011, 332:109-112.

26. Pörtner HO: Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. *Comp Biochem Physiol* 2002, 132A:739-761.

27. Farrell AP: Environment, antecedants and climate change: lessons from the study of temperature physiology and river migration of salmonids. *J Exp Biol* 2009, 212:3771-3780.

28. Cassinelli JD, Moffitt CM: Comparison of growth and stress in resident redband trout held in laboratory simulations of montane and desert summer temperature cycles. *Trans Am Fish Soc* 2010, 139:339-352.

29. Somero GN: The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine "winners" and "losers." *J Exp Biol* 2010, 213:912-920.

30. Fangue NA, Nann AF, Hofmeister M, Schulte PM: Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *J Exp Biol* 2006, 209:2859-2872.

31. Dong Y, Somero GN: Temperature adaptation of cytosolic malate dehydrogenases of limpets (genus *Lottia*): differences in stability and function due to minor changes in sequence correlate with biogeographic and vertical distributions. *J Exp Biol* 2008 212:169-177.

32. Miller KM, Li S, Kaukinen KH, Ginther N, Hammill E, Curtis JMR, Patterson DA, Sierocinski T, Donnison L, Pavlidis P, *et al.*: Genomic signatures predict migration and spawning failure in wild Canadian salmon. *Science* 2011, 331:214–217.

33. Fausch KD, Torgersen CE, Baxter CV, Li HW: Landscapes to riverscapes: bridging the gap between research and conservation of stream fishes. *BioScience* 2002, 52:483-498.

34. Ficke AD, Myrick CA, Hansen LJ: Potential impacts of global climate change of freshwater fisheries. *Rev Fish Biol Fisher* 2007, 17:581-613.

35. Gustafson RG, Waples RS, Myers JM, Weitkamp LA, Bryant GJ, Johnson OW, Hard JJ: Pacific salmon extinctions: quantifying lost and remaining diversity. *Conservation Biology* 2007, 21:1009-1020.

36. Perry AL, Low PJ, Ellis JR, Reynolds JD: Climate change and distribution shifts in marine fishes. *Science* 2005, 308:1912-1915.

37. Wenger SJ, Isaak DJ, Luce C, Neville HM, Fausch KD, Dunham JB, Dauwalter DC, Young MK, Elsner MM, Rieman BE, Hamlet AF, Williams JE: Flow regime, temperture, and biotic interactions drive differential declines of trout species under climate change. *Proc Nat Acad Sci* 2011, 108:14175-14180.

38. Isaak DJ, Rieman BE: Stream isotherm shifts from climate change and implications for distributions of ectothermic organisms. *Glob Change Biol* 2013, 19:742-751.

39. Allendorf FW, Hohenlohe PA, Luikart G: Genomics and the future of conservation genetics. Nature Reviews 2010, 11:697-709.

40. Cooke SJ, Sack L, Franklin CE, Farrell AP, Beardall J, Wikelski M, Chown SL: What is conservation physiology? Perspectives on an increasingly integrated and essential science. *Conserv Physiol* 2013, 1:1-23.

41. Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA: Genotyping-by-sequencing in ecological and conservation genomics. *Mol Ecol* 2013, 22:2841-2847.

42. Hoffmann AA, Sgro CM: Climate change and evolutionary adaptation. *Nature* 2011, 470:479–485.

43. Portner HO: Oxygen- and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. *J Exp Biol* 2010, 213:881-893.

Table 1. Summary data for redband trout samples including strain (LJ=Little Jacks Cr., K=Keithley Cr., LJxK=F1 crosses), temperature treatment (28°C treatment or 15°C control), sample day, sequencing reads (M=millions), and reference alignment statistics (transcriptome is abbreviated in column heading as Transc.).

Sample*	Strain	Temp	Day	Quality Reads (M)	Transcr. Match (M)	Mean Match(M)/ strain	Transc. Match (%)	Mean Match(%)/ strain
1	LJ	Trt-28C	1	28.1	4.6	7.0	16.2	23.1
2	LJ	Trt-28C	1	31.3	8.2		26.2	
3	LJ	Trt-28C	1	30.1	8.1		26.8	
4	LJ x K	Trt-28C	1	30.4	12.1	12.5	39.8	40.9
5	LJ x K	Trt-28C	1	27.6	11.6		42.1	
6	LJ x K	Trt-28C	1	33.6	13.7		40.9	
7	K	Trt-28C	1	32.4	13.1	12.1	40.4	40.3
8	K	Trt-28C	1	31.4	13.4		42.5	
9	K	Trt-28C	1	25.9	9.9		38.0	
10	LJ	Trt-28C	3	40.6	14.1	10.8	34.7	30.2
11	LJ	Trt-28C	3	34.8	9.8		28.3	
12	LJ	Trt-28C	3	30.5	8.4		27.7	
13	LJ x K	Trt-28C	3	31.3	9.1	9.0	29.0	26.1
14	LJ x K	Trt-28C	3	30.2	6.9		23.0	
15	LJ x K	Trt-28C	3	41.7	10.9		26.3	
16	K	Trt-28C	3	27.3	9.6	7.0	35.2	24.4
17	K	Trt-28C	3	28.3	7.6		27.0	
18	K	Trt-28C	3	34.3	3.8		11.0	
19	LJ	Trt-28C	7	29.3	12.4	9.1	42.4	29.2
20	LJ	Trt-28C	7	31.5	9.4		29.9	
21	LJ	Trt-28C	7	35.1	5.4		15.5	
22	LJ x K	Trt-28C	7	28.5	10.8	10.5	37.9	38.2
23	LJ x K	Trt-28C	7	26.4	9.9		37.6	
24	LJ x K	Trt-28C	7	28.0	10.9		39.1	
25	K	Trt-28C	7	37.4	6.8	5.2	18.2	16.6
26	K	Trt-28C	7	28.7	5.0		17.3	
27	K	Trt-28C	7	27.8	4.0		14.2	
28	LJ	Trt-28C	28	22.8	6.4	5.6	28.2	20.6
29	LJ	Trt-28C	28	31.2	7.5		24.0	
30	LJ	Trt-28C	28	29.6	2.8		9.6	
31	LJ x K	Trt-28C	28	43.0	0.5	3.7	1.1	14.9
32	LJ x K	Trt-28C	28	24.1	4.1		16.9	
33	LJ x K	Trt-28C	28	24.3	6.4		26.5	
34	K	Trt-28C	28	26.4	6.3	6.0	23.7	23.0
35	K	Trt-28C	28	25.4	7.3		28.7	

36	K	Trt-28C	28	26.8	4.5		16.6	
37	LJ	Con-15C	1	23.8	7.4	6.9	31.2	27.0
38	LJ	Con-15C	1	24.7	6.7		27.3	
39	LJ	Con-15C	1	29.4	6.6		22.5	
40	LJ x K	Con-15C	1	25.0	6.5	10.2	25.9	25.0
41	LJ x K	Con-15C	1	65.0	15.3		23.5	
42	LJ x K	Con-15C	1	34.2	8.7		25.6	
43	K	Con-15C	1	22.4	5.3	6.9	23.7	26.0
44	K	Con-15C	1	27.1	8.3		30.6	
45	K	Con-15C	1	29.6	7.0		23.7	
46	LJ	Con-15C	3	24.2	8.1	5.4	33.5	19.7
47	LJ	Con-15C	3	34.4	3.7		10.9	
48	LJ	Con-15C	3	30.0	4.4		14.8	
49	LJ x K	Con-15C	3	46.3	0.7	3.3	1.6	11.1
50	LJ x K	Con-15C	3	27.0	3.6		13.4	
51	LJ x K	Con-15C	3	30.6	5.6		18.4	
52	K	Con-15C	3	26.7	8.3	5.8	31.0	21.2
53	K	Con-15C	3	28.0	4.3		15.4	
54	K	Con-15C	3	28.5	4.9		17.1	
55	LJ	Con-15C	7	24.3	7.1	6.0	29.4	24.0
56	LJ	Con-15C	7	26.2	7.0		26.9	
57	LJ	Con-15C	7	25.0	4.0		15.9	
58	LJ x K	Con-15C	7	33.2	6.7	4.3	20.2	13.8
59	LJ x K	Con-15C	7	32.6	2.9		8.9	
60	LJ x K	Con-15C	7	27.3	3.4		12.4	
61	K	Con-15C	7	25.6	5.0	10.8	19.4	23.4
62	K	Con-15C	7	27.8	5.9		21.0	
63	K	Con-15C	7	72.7	21.7		29.9	
64	LJ	Con-15C	28	35.0	12.5	9.5	35.6	25.9
65	LJ	Con-15C	28	41.1	7.1		17.4	
66	LJ	Con-15C	28	36.1	8.9		24.6	
67	LJ x K	Con-15C	28	31.5	5.9	8.3	18.6	21.2
68	LJ x K	Con-15C	28	54.7	7.1		12.9	
69	LJ x K	Con-15C	28	37.3	11.9		32.1	
70	K	Con-15C	28	44.4	0.3	4.9	0.7	13.4
71	K	Con-15C	28	41.5	6.9		16.5	
72	K	Con-15C	28	32.3	7.4		22.9	
mean	--	--	--	31.96	7.53	7.5	24.1	24.1

*Each sample includes 3 pooled RNA samples from the same rearing tank

SECTION 2: Neutral and adaptive genetic variation of steelhead trout across highly variable landscapes

Introduction

Management strategies implemented for species conservation are highly contingent on a host of correlated life history and demographic information. In concert with these data, genetic structure is vital for characterizing population distinctions and limitations on productivity related to the decline of many species (e.g., conifer *Cathaya argyrophylla pineceae*, Ge et al. 1998; Chinese cobra *Naja atra*, Lin et al. 2012; gopher tortoise *Gopherus polyphemus*, Clostio et al. 2012; Chinook salmon, Moran et al. 2013). The genetic differentiation of populations across a species range is often determined on the basis of phylogenetic origins (e.g. Wagner et al. 2005), and historical and contemporary demography (e.g., Ruokonen et al. 2004). More recently, genetic variation has been viewed from the perspective of physical landscapes or environmental variation (Manel et al. 2003; Schoville et al. 2012). A landscape genetics approach reveals population variation relative to the influences or features in an organism's environment (Segelbacher 2010; Sepulveda-villet & Stepian 2012), including natural or human erected barriers and local climate. Most often it has been described on the basis of neutral divergence (Dionne et al. 2008; Narum et al. 2008), where restricted gene flow is explained in the context of a heterogeneous, patchwork environment (Latch et al. 2011).

Conservation units such as a *distinct population segments* (DPS) are established based on a core set of criteria including population ecology and viability, ancestry and descent, reproductive isolation, and local adaptation (Fraser and Bernatchez 2001; Fraser et al. 2011). Local adaptation may be inferred from neutral genetic structure coincident with habitat or life history variability (Olsen et al. 2011; Blankenship et al. 2012). However, direct evaluations of

non-neutral population differentiation (i.e. putatively adaptive divergence) are likely to reveal stronger, more correlative relationships (Limborg et al 2011). Nevertheless, conservation is not commonly informed by non-neutral variation, and inferences on adaptation based exclusively on neutral differentiation risk incorrectly identifying the underlying processes affecting population distinctions (Funk et al. 2012; Landguth & Balkenhol 2012).

Although full genome sequence data is typically not available for non-model species, evaluations of adaptive variation have recently been addressed using analyses of single nucleotide polymorphism (SNP) loci (Willing et al. 2010; Matala et al. 2011; Hohenlohe et al. 2010). Because SNP loci are commonly found within or adjacent to coding and regulatory regions of a genome, their allele frequencies may be influenced by selection (i.e. non-neutral). Techniques such as association testing and detection of outlier loci allow evaluation of differentiation that provides an improved understanding (over neutral loci) of the relationship between signatures of adaptive variation and the physical environment, even without direct interpretations of phenotypic variation, or interrogation of specific functional genes (Narum et al. 2010b; Matala et al. 2011; Ackerman et al. 2012a, Bourret et al. 2013).

Understanding the distribution of adaptive variation across landscapes will be crucial in establishing conservation priorities (see Crandall et al. 2000; Fraser & Bernatchez 2001), and for anticipating how populations might be affected by local and regional changes in climate (Hoderegger and Wagner 2008; Isaak et al. 2012a). The effects of global climate changes (e.g. rising temperatures) have increasingly altered habitats of myriad species, garnering the attention of a broad spectrum of researchers (Hickling et al. 2005; Milner et al. 2008; Winfield et al. 2010). Climate changes can prompt organisms to alter their behavior through range expansions (Loarie et al. 2009), or adapt over short time periods (rapid evolution; Hoffman and Sgro 2011;

Barrett et al. 2013). Some of the most profound examples are found among organisms limited by confined habitats such as those of fishes, whose habitats are prescribed by water routes (e.g., networks of streams and lakes). Because of this limitation they are particularly susceptible to environmental changes including water quality and temperature (Hari et al. 2006; Rieman et al. 2007; Wenger 2011). Fish adapt variably to thermal conditions in their migratory environment (thermal optimum for aerobic scope), and even populations within the same subbasin may be affected disproportionately by dramatic thermal shifts (Farrell et al. 2008).

In the Columbia River Basin (CRB), steelhead trout (*Oncorhynchus mykiss*) occur as two evolutionarily divergent lineages, delineated east (inland) and west (coastal) of the Cascade Mountain Crest. The inland redband trout (*O. m. gairdneri*) are typically a stream maturing, summer-run type, while coastal rainbow trout (*O. m. irideus*) are dominated by an ocean maturing, winter-run type (Busby et al. 1996; Behnke 2002; Currens et al. 2009; Blankenship et al. 2011). Some populations of *O. m. irideus* also have a summer-run life history, though not necessarily genetically differentiated from sympatric winter-run populations (Busby et al. 1996). Owing to persistent steelhead trout population declines throughout the region, managers have implemented extensive monitoring and evaluation efforts (Busby et al. 1996; Chilcote 1998; ICTRT 2003; Scott 2008; Fryer et al. 2012). Five Steelhead trout DPSs have been delineated within the CRB, and each is currently recognized for protection under the Endangered Species Act (ESA): the Upper Willamette River, Lower Columbia River, Middle Columbia River, Upper Columbia River, and the Snake River Basin (U.S. Office of the Federal Register 2006). Rivers in proximity to the Columbia River estuary lie within the Southwest Washington DPS. These conservation demarcations are largely contingent on adjacency of watersheds within stream networks, coupled with life history distinctions, and neutral genetic population structure. To date

there has been little to no direct interpretation of non-neutral variation for conservation assessment (Beacham 1999; ICTRT 2003; Good et al. 2005; USOFR 2006; Nielsen et al. 2009).

Studies that describe distinctions among particular steelhead trout populations are plentiful (e.g., Chilcote et al. 1986; Zimmerman and Reeves 2000; Hendry et al. 2002; Matala et al. 2008), and address some local conservation concerns. However, more extensive evaluations are necessary to accommodate broad conservation management priorities in a regional context (Beacham 1999; Winans et al. 2004; Currens et al. 2009; Blankenship et al. 2012). In this study, we investigate patterns of neutral genetic variation in contrast to non-neutral (putatively adaptive) genetic variation. We employed a multi-phased test approach to categorize non-neutrality (selection candidacy) of loci that was procedurally similar to several previous studies (Narum et al. 2010b; Hess and Narum 2011; Matala et al. 2011). Our primary objective was to provide an extensive characterization of genetic variation of steelhead trout throughout the entire Columbia River drainage to inform conservation management of this species. Study questions are threefold: 1) Is neutral population structure using SNPs consistent with previous studies that utilized other marker types [Winans et al. 2004; Currens 2009; Blankenship et al. 2012]?, 2) Is there significant evidence for candidate SNPs under selection and indications of putative adaptive variation?, and 3) How do patterns of neutral and non-neutral genetic variation compare and contrast among populations that occupy highly variable environments across the landscape. Lastly, we discuss how non-neutral differentiation in relation to climate change may improve our understanding of population viability, and promote informed conservation that will complement existing methods implemented for the practical management of many diverse and often imperiled species.

Methods

Sampling, genotyping and descriptive statistics

Our final data set consisted of 9,011 steelhead trout samples, representing 145 collections spanning the sampling years 1996-2011 (Table 1; Supplemental 1). Collections will hereafter be referred to as populations. The data set was primarily comprised of natural-origin populations (n=133) but some hatchery exceptions are included (n=12). Both the coastal and inland lineages were represented but in skewed numbers of populations: coastal (n=24), and inland (n=121; Supplemental 1; Figure 1). Sample sizes genotyped per population and by major population group (MPG) had minimum numbers ranging from n=18 to n=90, and maximums ranging from n=30 to n=164; Table 1).Genomic DNA was extracted from fin or opercle tissues of juvenile and adult fish preserved dry on Whatman paper (see; LaHood 2008) or stored in individual vials containing 100% non-denatured ethanol. For DNA extraction we used a standard Qiagen® DNeasy™ protocol, or Nexttec™ Genomic DNA Isolation Kits from XpressBio (Thurmont, Maryland) following the manufacturer's standard protocol.

A total of 191 SNP loci were genotyped with Taqman assays (Applied Biosystems. Our locus panels were comprised of SNPs developed by multiple sources (Supplemental 2a). All loci were ascertained from a broad coast-wide sample of populations, including Alaska, Washington, British Columbia, California, and Russia. Most SNPs were developed in a process of mining expressed sequence tags (ESTs) from GeneBank (https://www.ncbi.nlm.nih.gov/genbank/), followed by sequencing via primers developed for EST sequence flanking regions. Generally, the coding or non-coding nature of specific ESTs, and associated functions were unknown (see references for additional information, Supplemental 2). Forward and reverse primer/probe

sequences for Taqman assays are available in Ackerman et al. (2012b). Polymerase chain

reaction (PCR) amplification of loci followed the protocol of Hess et al. (2012), and included an

initial pre-amplification step. Successful genotyping for a given sample was defined

proportionally as less than 10% missing data (i.e. fewer than 19 of 191 SNP genotypes per

individual). Three species specific SNP markers (*Ocl_gshpx-357*, *Omy_myclarp404-111*, and

Omy_Omyclmk438-96) were used to screen for species ID and hybridization between *O. mykiss*

and *O. clarkii* subspecies (Hess et al. 2012). All individuals identified as hybrids and congeners

(n=15 coastal, n=79 inland) were subsequently removed from the data set prior to analyses.

Following this screening exercise, the three species-diagnostic loci were removed from the data

set.

Ten SNPs in our pared panel of 188 SNPs have been previously identified as *a-priori*

candidate loci for selection. Specifically, two loci are putatively associated with thermal stress-

induced mortality (Narum et al. 2013): *Omy_hsp47-86*, and *Omy_OmyP9-180*. Five SNP loci

(*Omy_aldB-165*, *Omy_gdh-271*, *Omy_Ogo4-212*, *Omy_stat3-273*, and *Omy_tlr5-205*) have been

previously identified as associated with temperature variation in desert vs. montane

environments, and one locus (*Omy_hsf2-146*) was putatively associated with precipitation in the

same study (Narum et al. 2010b). Additionally, two loci have been shown to differentiate

anadromous and resident life history types (Narum et al. 2011): *Omy_ndk-152* and *Omy_LDHB-

2_i6*. In the following sections these ten loci are referred to as having "precedence" in

association tests (Supplemental 2b).

Locus specific allele frequencies (i.e. minor allele frequency; MAF), observed

heterozygosity (H$_o$) and F_{ST} were generated with the program GenAlEx version 6.2 (Peakall and

Smouse 2006).   Pairwise population F_{ST} was calculated in GENEPOP v. 3.3 (Raymond and

Rousset 1995) and significant among-group variation was determined at P<0.01 using

ARLEQUIN version 3.5 (Excoffier et al. 2005). Pairwise stream distances between population

pairs were calculated using ArcMap and a GIS application developed by D. Graves (CRITFC),

available at: http://maps.critfc.org/file_download/StreamDistanceApplication.zip. Mantel tests

of isolation by distance (IBD) were evaluated in GenAlEx v.6.2 using matrices of pairwise F_{ST}

and pairwise stream distance. The Markov Chain Monte Carlo (MCMC) approximation of

Fisher's exact test implemented in GENEPOP v. 3.3 (1000 batches with 1000 iterations) was

used to test for deviations from Hardy-Weinberg equilibrium (HWE) expectations, evaluated

across 188 SNP loci and 145 populations. Linkage disequilibrium was tested for all pairs of loci

among populations using a simulated exact test in GENEPOP. For all pairs of loci with

significant non-random association (linkage), the locus with the lower MAF was excluded from

further analyses. Statistical significance (   ) was adjusted for the number of simultaneous tests

(initial   = 0.05) for both HWE and linkage tests via the B-Y FDR method (Benjamini and

Yekutieli 2001) to reduce false positive tests.

Population structure within and among lineages

Throughout our methods and analytical approach the coastal and inland lineages were

evaluated separately. First, we verified the resolving power of our SNP markers to discretely

differentiate the two steelhead lineages that have been characterized in previous studies based on

other genetic markers (e.g., allozymes, Currens et al. 2009; microsatellites Blankenship et al.

2011), while confirming lineage-of-origin for each population. This test was conducted

irrespective of classification of loci as neutral or non-neutral. The program STRUCTURE

version 2.3.4 (Pritchard et al. 2000) was used to estimate the mean admixture proportions (*Q*), or

fractional group fidelity among individuals in each of 145 populations, setting *k*=2 (two inferred

groups; coastal and inland). Default settings were used for the ancestry model (i.e. admixture) and frequency model (i.e. correlated allele frequencies), with a burn-in of 50,000 and 250,000 MCMC repetitions. Locus specific and global pairwise F_{ST} (of Weir and Cockerham 1984) were calculated within each lineage using GENEPOP. A multivariate principle coordinates analysis (PCoA) was performed in GenAlEx version 6.2 based on matrices of pairwise F_{ST} to quantify amounts of variation.

Identifying putatively neutral loci

In classifying the nature of SNP loci we define three possible outcomes or distinctions. A locus may be 1 - unaffected by selection (neutral), 2 - directly under selection (affecting one or more traits under selection), or 3 - indirectly affected by the process of selection (selection at a linked locus). In all subsequent discussion, the term "non-neutral" variation will be used to represent the latter two distinctions. We used an outlier approach based on simulation methods initially proposed by Beaumont and Nichols (1996) to identify outlier SNP loci putatively under selection. We conducted outlier tests as implemented in ARLEQUIN version 3.5 (Excoffier et al. 2005) to test for excessively higher or lower F_{ST} than would be expected under the assumptions of neutrality (Beaumont and Nichols 1996; Beaumont and Balding 2004). Tests were performed using 100 simulated demes and 50,000 coalescent simulations to generate a null distribution under neutral expectations around observed F_{ST} values (with confidence intervals). Due to the potential high error rate of the hierarchical island model (Narum and Hess 2011), a finite island model was assumed for outlier tests. Average locus heterozygosity versus F_{ST} was plotted to represent the 1% and 99% quantiles, and loci lying below or above these quantiles were outliers putatively under balancing or directional selection, respectively.

Outlier loci can be neutral or non-neutral and therefore relying solely on outlier tests to characterize the selection candidacy of loci risks a resulting high rate of false positives. For example, populations may be highly differentiated because of demography, or skewed patterns of isolation-by-distance (Akey 2009; Hermisson 2009; Narum and Hess 2011). In fractal landscapes such as stream networks, patterns of genetic structure can be coincident with patterns or complexity of stream branching, likely resulting in elevated variance around neutral F_{ST} (i.e. false positive outliers; Fourcade et al. 2013). For our exploration of non-neutral differentiation, outlier methods were used in combination with regression models with control variables to test environmental associations. This reduces the risk of false positive conclusions (e.g., Matala et al. 2011), and aids in identifying the influences of environmental or habitat heterogeneity on the spatial distribution of genetic diversity (inferred selection gradients).

Defining physical and control variables in an association test framework

Physical habitat variables used to gauge environmental heterogeneity were chosen to represent regional and local climate regimes. Latitude and longitude coordinates for each population were obtained from field data or were estimated using ArcMap version 10 (copyright © 2010 ESRI), and were used to gather all physical variable measurements (Supplemental 3a & 3b). Migration distance from the Columbia River estuary and pairwise stream distances (kilometers) were calculated using a GIS application developed by D. Graves (CRITFC), available at: http://maps.critfc.org/file_download/StreamDistanceApplication.zip. Elevation was obtained using Google Earth (Image © 2012 TerraMetrics) and verified for accuracy in ArcMap version 10. Monthly averages for temperature and total precipitation (rain and melted snow) measurements were generated using Parameter-elevation Regressions on Independent Slopes Model (PRISM) of the Oregon Climate Service; http://www.prism.oregonstate.edu/. Monthly

average maximum and minimum air temperatures were simulated at 800-meter cell resolution

from a model based on climate normals from a 30 year period (1971-2000) in PRISM. Water

temperature readings were unavailable and therefore we used air temperature readings as a proxy

for in-stream temperature. In the Pacific Northwest region, stream temperature trends closely

with air temperature, particularly on temporal scales (Isaak et al. 2012b). However, changes in

stream temperatures occur more slowly and are affected by groundwater and riparian buffering,

glaciation, and complex topography. In total, five physical variables were used to characterize

local and regional climate: precipitation, temperature, elevation, geographic coordinates, and

migration distance. We expected autocorrelation between all five variables. The precipitation

and temperature variables we evaluated on the basis of annual and seasonal measurements.

Geographic coordinates and migration distance add directional elements related to regional

weather patterns.

The highly dependent nature of F_{ST} on demographic history may confound the ability to

accurately identify selection candidate loci (Beaumont 2005). When isolating forces appear

correlated with variation across the landscape (Narum and Hess 2011), it can be difficult to

confidently differentiate demographic influences from non-neutral (putatively adaptive)

associations. Prior to conducting association tests, we assembled panels of putatively neutral

SNP loci (defined by exclusion of outliers) for each lineage, and as a conservative measure we

established control variables for underlying neutral population structure. Using STRUCTURE

version 2.3.4 (Pritchard et al. 2000) we evaluated k ranging from 1-8 clusters for each lineage,

and the most likely number of distinct populations (k) was selected based on five iterations for

each potential k value. We used program default settings for the Monte Carlo Markov Chain

procedure, with 30,000 burn-in followed by 150,000 MCMC repetitions. The (Δk) statistic

derived from the second order rate of change of the likelihood function (Evanno et al. 2005) provided an improved estimate of the mode of true (*k*), and the mean Q values for each population were established using the full search algorithm in CLUMPP (Jakobsson and Rosenberg 2007). Secondly, a pairwise population F_{ST} matrix was generated in GENEPOP to conduct PCoA in GenAlEx version 6.2. The resulting (*Q*) inferred admixture proportions from STRUCTURE and the first three Eigen vectors (EV) from PCoA analysis were used to control for underlying neutral population differentiation in subsequent association tests.

Identifying candidate loci: Association with Environment

We used two regression methods to identify potential candidate markers in association with the previously described environmental or climate related variables. First, population MAF at each locus was plotted against each of five physical variables (Supplemental 2a) in univariate linear regression analysis. The *P*-values for the correlation coefficient were calculated in Microsoft Excel. Statistical significance was adjusted for the number of simultaneous tests as a conservative measure using a Bonferroni correction to exclude false positive results (Rice 1989). Eighteen populations from the genetically distinct Clearwater River are located in a geographic region characterized by high precipitation. All 18 ranked in the top 20 for greatest amount of spring and summer precipitation among 121 inland populations (variable non-independence). In initial univariate regression analyses, 86 loci (48%) showed significant correlation with either spring or summer precipitation. Therefore, as a precautionary measure to guard against what appeared to be a spurious geographical influence, regression coefficients were recalculated in the absence of the Clearwater River group. Then, of the original 86 loci only those that maintained a significant correlation with precipitation were considered "true" correlations.

Second, we used DISTLM *forward* (McArdle and Anderson 2001) to perform

multivariate multiple regression with permutation tests on locus-specific pairwise F_{ST} distance

matrices versus pairwise matrices of values for physical variables (Carmichael et al. 2007; Olsen

et al. 2010; Hess et al 2011; Matala et al. 2011). Conditional tests that permute residuals under a

reduced model (Anderson 2003) were performed in DISTLM *forward* using a stepwise forward

selection procedure that fits individual variables or sets of variables sequentially in the linear

model. Once the most informative variable or set of variables (explaining greatest amount of

variation) is established, the remaining variables are fit to the model while those selected in

previous steps are held as constants. Specific associations were determined on the basis of

highest rank (*P*<5%) variable for a given locus, and the method accounts for correlations that are

likely present between the variables used to characterize climates (Supplemental 3 & 4). The

temperature, precipitation, and neutral-control (*Q* and EV) variables were evaluated in respective

"sets" as demonstrated by Anderson (2004), where seasonal averages for the former two were

defined as: winter (Jan.-Mar.), spring (Apr.-June), summer (July-Sept.), and fall (Oct.-Dec.).

In summary, the first two criteria used to flag loci for further consideration as candidates

in our multi-test process of categorization were: 1) significant F_{ST} outliers at a 99% confidence

threshold or 2) loci with precedence of association [see Supplemental 2b]. All 180 loci were

tested for correlation based on linear regression (criterion #3). Lastly, all loci flagged as

candidates based on the first three criteria were scrutinized on the basis of multivariate multiple

regression with permutation tests (criterion #4). Note that as the final discriminatory step, when

multivariate regression tests ranked a neutral control variable (*Q* or EV) as highest in

significance among all tested variables, the locus in question was precluded from categorization

as non-neutral (selection candidacy) regardless of results based on the other three criteria (e.g.

linear correlation with environment). This approach resulted in three sub-panels of SNP loci for genetic analysis: candidate loci – those meeting stipulations outlined above, neutral loci – those showing strongest correlation with neutral structure, and "ambiguous" loci. The latter category of SNP was primarily comprised of F_{ST} outlier loci or loci having precedence of association which ultimately failed to reach non-neutral status based on regression analysis. Ambiguous loci (by definition) could not be confidently characterized as neutral or non-neutral given conflicting test results and were therefore excluded from subsequent evaluations of genetic differentiation. Candidate loci putatively under selection were subsequently used to evaluate non-neutral or adaptive variation. Putatively neutral loci were used to evaluate underlying neutral (demographic) population structure.

Comparing neutral and putatively adaptive variation

Using the newly established neutral and candidate SNP panels, comparisons were drawn based on multiple analyses to show the corresponding amount of variation described by both neutral and non-neutral variation (Nosil et al. 2007). Nonparametric Mantel tests for isolation as a function of environment were conducted using 9999 permutations of pairwise matrices of F_{ST} (within lineages) against absolute pairwise difference in values for environmental variables. Nei's standard genetic distance (Nei 1972) was calculated for each lineage, and distance was displayed in the topology of an un-rooted neighbor-joining (NJ) tree using the analysis program PHYLIP version 3.68 (Felsenstein 2008). The SEQBOOT option was implemented to generate 1000 simulated data sets, and a consensus topology with bootstrap support was generated using the CONSENSE option. The program TreeView version 1.6.6 (Page 1996) was used to graphically display the trees. Different trees were generated using either neutral or candidate SNP panels for each lineage.

A ranking method was employed based on averages for five climate variables to compare climate similarities that occurred in the clustering of populations within tree topologies. Specifically, populations were ranked in descending order (warmest to coolest) using corresponding mean maximum temperatures for both spring and summer independently. Elevation was ranked in ascending order assuming lowest elevation equals warmest climate. Lastly, populations were ranked in ascending order for mean precipitation (least equal to driest/warmest climate) in both spring and summer independently. Following independent ranking of populations for the five variables, the average rank across all five was used to order populations from most hot and dry to most cold and wet, and to compare relative climates in the tree topologies.

Results

Descriptive statistics and population differentiation

Data quality analyses indicated only minor issues concerning locus scoring accuracy, non-random association of loci, and population admixture. We observed departures from expected genotypic proportions in 208 out of 27,260 tests (188 loci x 145 populations) at an adjusted significance threshold of P=0.0046. Generally, the HWE deviations were not specific to any population or locus, spanning 100 of 145 populations, and 109 of 188 loci. Exceptions occurred in both Abernathy Creek (Ref. #12) and Canyon Creek (Ref. #9), each with 10 population-specific departures. There were also 12 HWE departures at locus *OMS00087*, which was therefore removed from all subsequent analyses. Tests for linkage disequilibrium revealed five pairs and one trio of loci that remained significantly out of equilibrium in at least 10% of

populations after adjustment for multiple tests (*P*-value <0.0001). Linked SNP pairs were: (*OMS00133* and *Omy_rapd-167*), (*Omy_CRBF1-1* and *Omy_crb-106*), (*Omy_Il-1b_.028* and *Omy_Il1b-198*), (*Omy_SECC22b-88* and *OMS00169*), (*Omy_ndk-152* and *Omy_u09-52.284*), and the trio (*Omy_GHSR-121*, *OMS00176* and *Omy_mapK3-103*). In each pair, only the locus with the highest minor allele frequency was retained in the data set (Supplemental 2a).

Following paring of 8 loci for HWE and linkage disequilibrium, the final data set included 180 loci for use in subsequent analyses. No SNP locus exhibited fixed allele frequencies, and variability ranged widely both within and between steelhead trout lineages. Among 24 coastal lineage populations, the mean observed heterozygosity ranged from H_o=0.002 at locus Omy_pad-196 to H_o =0.531 at locus OMS00101 (overall mean H_o =0.315). Among 121 inland lineage populations the range was H_o=0.027 at locus Omy_nach-200 to H_o =0.513 at locus Omy_IL17-185 (overall mean H_o =0.302; Supplemental 2a). Among coastal lineage populations we observed a mean locus MAF of 0.246 ranging from 0.001 at *Omy_impa1-55* to 0.498 at *Omy_arp-630*. The mean MAF across inland lineage populations was 0.236, ranging from 0.024 at *Omy_nach-200* to 0.497 at *OMS00070*. Generally large MAFs among populations within lineages are indicative of highly differentiated populations. The overall mean pairwise F_{ST} between all coastal populations was 0.044. Population specific mean values among 24 populations ranged from pairwise F_{ST} =0.029 (ref.#'s 17) to pairwise F_{ST} =0.079 (ref.#'s 11) . With the exception of one pair of populations (ref.#'s 3 & 4; Eagle Creek, North Fork Eagle Creek) all comparisons indicated significant among-group variation (P<0.001). The overall mean pairwise F_{ST} between all inland populations was 0.045. Population specific mean values among 121 populations ranged from pairwise F_{ST} =0.025 (ref.#'s 33) to pairwise F_{ST} =0.148 (ref.#'s 77), and all pairwise population comparisons indicated significant among-group variation (P<0.001).

Significant isolation by distance (IBD) was observed in both lineages: coastal (R^2=0.245, P<0.0001) and inland (R^2=0.083, P<0.0001).

Differentiating populations by lineage

When all coastal lineage populations were combined and all inland lineage populations were combined to form two groups (corresponding to lineage), we observed distinct allele frequency differences over the panel of 180 loci. Allele frequencies ranged from a low of 0.002 at *Omy_104519-624* to a high of 0.719 at *Omy_ndk-152*. At 41 different (bi-allelic) loci the minor allele in one lineage was the opposite (major allele) for the other lineage. The overall mean pairwise F_{ST} for comparisons between coastal and inland populations was 0.145. Population-specific means for inter-lineage comparisons ranged from pairwise F_{ST} =0.053 (ref.#'s 25: Bowman Creek) in the inland lineage, to pairwise F_{ST} =0.228 (ref.#'s 33; Upper Trout Creek) also an inland lineage population (data not shown). With the exception of 15 populations along the crest of the Cascade Mountains, the 145 populations in our analyses formed defined clusters according to lineage, where the first PCoA plot axis separated lineages and explained 63.4% of the total variation in the data. In Bayesian clustering analyses, the mean admixture proportion (*Q*) for two inferred populations was 95.3% in *Q*1 for coastal populations, and conversely 93.8% for *Q*2 among inland populations (Figure 2). Mean values would likely have been higher, but populations adjacent to the crest of the Cascade Mountains (demarcating range limits of coastal and inland types) appeared admixed between lineages to varying degrees, substantially lowering inferred group fidelity. The genetic characterization of the Big White Salmon River population (Figure 1; c) was significantly more similar to the coastal lineage despite its location among the middle Columbia River DPS, and was therefore evaluated throughout these analyses as a coastal population.

Controlling for underlying neutral differentiation in association testing

Outlier tests revealed eight loci putatively under directional selection in the coastal lineage, and ten in the interior lineage (Supplemental 2b & 4). No influence of balancing selection was observed. The remaining putative neutral loci in the inland lineage (n=170) and coastal lineage (n=172) were used in PCoA and Bayesian cluster analyses to establish neutral control variables. For both lineages the most likely number of inferred populations was k=2. However, this was likely a result of the relatively deep divergence between the Willamette River DPS and remaining lower Columbia River populations in the coastal lineage (ICTRT 2003), and similar divergence between the genetically distinct Klickitat River subbasin and remaining inland populations. To evaluate structure at a finer scale we chose the appropriate number of inferred groups from the next peak in the second order rate of change (Δk) for $\log_e[\Pr\{K\}]$, which occurred at k=6 ($Q1$-$Q6$) within each lineage independently (Supplemental 3a).

For the coastal lineage, mean Q partially distinguished summer-run populations, a less common life history type among this lineage ($Q5$=65%), two groups in the upper Willamette DPS ($Q2$=58% and $Q4$=61%), the Washington coast population from the Quinault River ($Q1$=89%), and the Big White Salmon population on the Cascade Crest ($Q6$=81%). Most of the remaining lower Columbia populations had greatest proportions in $Q3$, ranging from 34% to 77%. For the inland lineage, the six inferred populations partially distinguished major subbasins, where the middle and south forks of the Salmon River were both dominant in $Q1$ (74%), the south and middle forks of the Clearwater River in $Q4$ (53%) and $Q3$ (76%) respectively, the Klickitat River in $Q2$ (68%), and the upper Salmon River in $Q5$ (56%). Regions with highest mean admixture proportion in $Q6$ included the majority from the upper and middle Columbia, and the Grande Ronde and Imnaha rivers in the Snake River Basin. Most individual populations

within these regions did not exhibit definitive or strong fidelity in $Q6$ (ranging 27% to 68%, with

a mean of 44%), and the highest admixture proportion by watershed occurred in the Yakima and

John Day Rivers, with mean $Q6$ of 59% and 54% respectively (Supplemental 3a). Differentiation

was slightly higher within the inland lineage, with a larger range of pairwise population F_{ST}

(0.0002 to 0.1889, mean 0.0421) compared to values within the coastal lineage (range 0.0001 to

0.1098, mean 0.0413). Results of PCoA based on pairwise F_{ST} corroborate primary distinctions

revealed from STRUCTURE analyses. Among the inland lineage, observed distinctions

coincide well with several of the Snake River DPS major population groups (MPGs).

Association tests to identify non-neutral loci: putatively adaptive variation

The adjusted probability threshold for determining significant association using linear

regression was $P<0.00029$. On the basis of linear regression alone, we identified 12 loci in the

coastal lineage that were significantly correlated with one or more environmental variables

(Table 2; Supplemental 2b). In the inland lineage 59 loci were identified as significantly

correlated, including 17 loci for spring precipitation, and 21 loci for summer precipitation. From

combined results using outlier tests, association precedence, and linear regression, we ultimately

flagged 26 loci in the coastal lineage and 62 loci in the inland lineage for further examination of

environmental correlation. Following the final ranking phase based on multivariate regression

the pared locus classifications included nine candidate loci putatively under selection in the

coastal lineage and 22 candidates in the inland lineage; two candidate loci were significant

within both lineages (Table 2). All remaining loci were considered either neutral or ambiguous

(Supplemental 2b). The group of loci ultimately classified as ambiguous were typically

comprised of outlier loci for which the highest ranked variable was one of our control variables

for neutral differentiation (Q or EV; Supplemental 3a). The SNPs deemed ambiguous may

46

indeed be under selection but we failed to identify associations with the suite of environmental variables tested here.

Of the eight SNP loci showing precedence as candidates for local climate from previous studies, seven exhibited significant association with climate variables in at least one lineage in our broader scale evaluation across many populations. Two loci previously identified as associated with survival under thermal stress (Omy_hsp47-86 and Omy-P9-180; Narum et al. 2013), were significantly associated with either temperature or precipitation (Figure 3). Two loci previously identified as candidates for temperature (*Omy_stat3-273 and Omy-gdh-271;* Narum et al. 2010b) were found to be highly correlated with precipitation in at least one lineage (Table 2; Supplemental 4b), but not for temperature. Previous associations of locus *Omy_hsf2-146* with precipitation, and locus *Omy_aldB-165* with temperature (Narum et al. 2010b) were corroborated within the coastal lineage. Finally, locus *Omy_tlr5-205* previously associated with temperature (Narum et al. 2010b) was significant for temperature and precipitation in the inland lineage. Notably, two of these loci (*Omy-P9-180, Omy_stat3-273*) were candidates for climate association in both lineages. Additional novel candidate loci were detected that have not previously been identified as loci putatively under selection.

The two loci that were previously determined to be associated with differentiating resident from anadromous life history forms (*Omy_ndk-152* and *Omy_LDHB-2_i6*; Narum et al. 2011) could not be directly tested for association with life history forms since most populations in our analyses were field identified as anadromous (e.g., juvenile smolt or adult steelhead phenotypes). However, *Omy_ndk-152* was a significant F_{ST} outlier. Two loci with precedence of putative association with climate or life history had no confirmed associations for any habitat

variable in our analyses and were deemed ambiguous since they may be candidates at smaller, more local scales.

Among coastal lineage populations, precipitation and distance from the ocean (i.e. migration distance, and lat/long coordinates) were equally the most commonly correlated environmental factors. To clarify, the common point of origin for all measured migration distances was the Columbia River estuary, therefore distance associations were not necessarily an example of IBD gene flow which relates to direct distance between populations. In the inland lineage environmental correlations were dominated by precipitation, then temperature. Specifically, spring, summer and total annual precipitation were the physical variables most often associated with genetic variation, and several significantly correlated loci spanned both lineages (Figure 3). In some cases, the ranking of variables in multivariate regression was inconsistent between tests of individual variables versus sets of variables. For example, while an individual variable (e.g., $Q1$, summer precipitation) may have ranked highest for a given locus, the corresponding variable sets (e.g., $Q1$-$Q6$, total precipitation) may have ranked low for that same locus (Supplemental 4a & 4b). An explanation for this outcome, centered on differential environmental influences among life history stages, follows in the discussion.

Comparing neutral and non-neutral differentiation

The NJ tree topologies based on neutral SNP panels generally showed genetic distance relationships among populations that accurately aligned with the five distinct population segments delineated under the ESA (Figures 4a & 5a). However, in the coastal lineage the Clackamas River populations (Lower Columbia River DPS) grouped closely with populations in the upper Willamette River DPS, while populations in the upper Willamette River west side

48

tributaries were distinctly partitioned from upper Willamette east slope populations (Figure 4a). The known summer-run populations in the coastal lineage cluster together with significant bootstrap support, rather than clustering with winter-run populations from the same tributaries (i.e. Kalama and Hood rivers). Middle and Upper Columbia MPGs and five MPGs in the Snake River (Ford et al. 2011) are well differentiated within the inland lineage, although bootstrap support was minimal in some instances. Finer scale definition observed in the Salmon and Clearwater rivers indicates significant genetic distinctions between middle and south fork populations from both subbasins (within MPGs), and between each of those groups and corresponding populations in the lower sections of both subbasins (Figure 5a). These within-watershed distinctions in both the Clearwater River and Salmon River subbasins are in agreement with previous reports (Nielsen et al. 2009), but the level of resolution that differentiates the Clearwater River subbasin from the Salmon River subbasin has not been previously reported. In the tree topology for neutral structure in the inland lineage, populations within MPGs or subbasins were also frequently characterized by climate similarity. However, deviations from ESA- or regionally-based (e.g., distance) clustering were rare regardless of differences or similarities in climate.

The NJ tree topologies were based on respective panels of total numbers of candidate SNPs in the coastal (n=9 loci) and inland (n=22 loci) lineages. Trees presented population clustering patterns that often did not correspond with delineations of DPS or MPG (Figures 4b & 5b), albeit with limited bootstrap support; presumably as a function of topologies based on small numbers of loci. Climate based clustering patterns were more apparent in the inland lineage, where several of the relationships depicted in the non-neutral tree topology conformed to warm or cool climate similarity irrespective of subbasin or MPG distinction. For example, the neutrally

49

distinct Klickitat River and Yakima River groups were each split to show some association to climate, most Clearwater and Salmon river groups with similar climates were combined on same primary branch of the tree, and the five lower Clearwater River populations were differentiated by climate distinctions. Overall, NJ results based on non-neutral SNPs for the coastal lineage make it difficult to discern any basis for population similarities within branching patterns. In particular, climate rankings for coastal populations were non-informative for understanding genetic distance relationships associated with physical variables (Figure 4a; climate ranks not shown).

Allele frequencies were most commonly correlated with the precipitation variables in association tests. This was further verified with Mantel tests of isolation-by-precipitation (IBP). Tests were based on a five-locus subset of nine candidate loci in the coastal lineage and a 15-locus subset of 22 candidate loci in the inland lineage; subsets were chosen based on significant associated with precipitation specifically (Table 2; supplemental 4a & 4b). For the coastal lineage, a test based on all five loci indicated significant IBP for mean annual precipitation (R^2=0.235; P=0.0009), mean spring precipitation (R^2=0.201; P=0.0006) and mean summer precipitation (R^2=0.296; P=0.0004). The panel of 158 neutral loci for the coastal lineage was also tested to account for potentially confounding neutral patterns of IBP, but none were observed. Mantel tests of IBP based on 15 precipitation-associated loci in the inland lineage (pairwise comparisons between 121 populations) indicated significant IBP for mean spring precipitation (R^2=0.039; P=0.0001), and for mean summer precipitation (R^2=0.037; P=0.0001); no correlations were observed between seasonal precipitation and the panel of inland lineage neutral SNPs (n=146).

Discussion

This study demonstrates patterns of neutral and non-neutral variation in *O. mykiss* at a broad geographic scale that will be a valuable contribution to improved conservation management of this species in the Columbia River Basin. Neutral structure was complementary to preceding studies (e.g., Winans et al. 2004; Currens et al. 2009; Blankenship et al. 2011), and confirms the existence of two deeply divergent steelhead trout lineages (i.e. coastal and inland) across the species range, along with finer scale population structure. We identified neutral divergence among steelhead trout populations that coincides substantially with current DPS delineations among lineages, as well as major population groups (ICTRT 2003; Good et al. 2005; Ford 2011) demarcated within the Snake River Basin (five MPGs), Middle Columbia River (four MPGs) and Upper Columbia River (one MPG). However, using our SNP panel, fine-scale resolution of neutral differentiation was detected in the Snake River DPS at a level previously unreported based on other marker types. Similar to other studies (e.g. Nielsen et al. 2009; Campbell et al. 2012) we found no evidence for multiple evolutionary lineages in the Snake River, but local environments may influence their differentiation. For reference, six Snake River hatchery stocks were genetically similar to their natural-origin counterparts that distinguish the Snake River DPS. In the coastal lineage, neutral differentiation among populations from east and west side tributaries in the upper Willamette River does not appear to fit DPS unit distinctions and may warrant further investigation to define conservation boundaries. From our analyses, the Big White Salmon river population is currently the only population within the Middle Columbia River DPS that is more consistent with a coastal lineage origin, suggesting further evaluation of its classification may be necessary.

Additionally, we show clear evidence for non-neutral (putatively adaptive) variation that is significantly associated with climate in the region. Specifically, several candidate markers were primarily associated with precipitation and temperature. This study demonstrates that candidate markers can be applied at broad geographic scales to describe the extent of potential local adaptations across highly variable climates. To evaluate non-neutral differentiation, our designation of candidate SNP loci was applied conservatively to reduce false positive results. Conclusions of climate association were based on a large and diverse number of populations, and supported by a framework of multiple test criteria and strict likelihood thresholds (Balkenhol et al. 2009; Schoville et al. 2012). De Mita et al. (2013) suggest using multiple robust methods, and emphasize numbers of populations over numbers of individuals per population for improved confidence in determining selection candidacy of loci. A greater number of candidate loci were discovered in the inland lineage, represented by a larger number of populations than were observed in the coastal lineage. In addition, the interior region of the CRB is larger, characterized by a highly variable environment relative to the coastal environment (greater range of wet/dry and warm/cool conditions). Several loci were identified in association with tested environment or landscape variables, but precipitation and temperature proved to be the most common (# loci) and strongest factors in non-neutral population differentiation that spanned both lineages (e.g., at SNP *Omy_stat3-273*, *Omy_OmyP9-180*).

Although our association tests indicated relationships between environmental variation and genetic heterogeneity (i.e. allele frequency variation) it is challenging to decipher the biological relevance of those correlations. Climate variables such as temperature may affect behaviors and phenotypes alike (Perry et al. 2001; Zydlewski et al. 2005), sometimes relatively rapidly in response to perturbations (Kovach et al. 2012). Previous landscape genetics studies in

salmonids have shown significant allele frequency correlation with precipitation that have been described as neutral influences on population structure (Narum et al. 2008; Blankenship et al. 2011; Olsen et al. 2011). Olsen et al. (2011) for example, present a compelling discussion on possible correlations between precipitation and gene flow. This may occur if increased flooding results in decreased stream stability, which in turn may affect fish dispersal. Alternatively, the similarity of outcomes across study locales may suggest that the distribution of precipitation across the landscape elicits, or is indicative of common adaptive responses. For example, streambed scour related to rain-on-snow events may have lesser impact on fish that adapt by burying eggs at deeper depth (Goode et al. 2013). Thus climatic variables such as precipitation may impart either "neutral" landscape effects, "non-neutral" (putatively adaptive) landscape effects, or both. In either case, one could argue that adaptation and selection play a key role in shaping the genetic landscape, which is more apt to be revealed through non-neutral genetic variation (Limborg et al. 2011).

The genetic population structure we observed differed markedly depending on whether our evaluations were based on neutral or non-neutral differentiation. Neutral differentiation generally reflected a pattern of distance-restricted gene flow, and in the context of demographic factors, population clustering was relatively intuitive within and among regions (clustering by subbasins, DPS, etc.). In contrast, population clustering patterns identified using candidate loci (non-neutral differentiation) were presumably based upon environmental variation, and frequently did not align with current steelhead trout DPS delineations. Moreover, clustering similarities based on non-neutral variation did not necessarily coincide with geographic proximity, nor were patterns among populations always transparent in regard to biology (e.g., coastal lineage migration timing). Thus the juxtaposition of genetic signals show that neutrally

dissimilar populations may exhibit non-neutral similarity (and vice versa) related to environment, and irrespective of geographic distance. Compared to neutral variation, candidate loci revealed some novel distinctions between populations, presumably reflective of environmental differences between regions and locales, particularly for the inland lineage. For the coastal lineage, less variable environments may produce moderate selective pressure for local adaptation among surveyed portions of the species range. More extreme environments in the southern portion of the coastal lineage range (e.g. the Sacramento River system) might be expected to provide stronger selective pressure.

The relative influences of climate or environment on genetic variation (e.g., putative adaptive responses) may occur during particular life history stages of an organism, which can be difficult to discern. Temperature for example has been shown to have variable effect on different life stages of fish (Fowler et al. 2009). In our multivariate regression analysis we noted inconsistencies between tests on individual physical variables and corresponding sets of seasonal variables. Seasonal environmental variation may impart a disproportionate selective influence coincident with age related behaviors (e.g., emergence time or outmigration time; Coleman and Fausch 2007). Our results of non-neutral differentiation indicate that precipitation during specific juvenile rearing or adult spawning periods may be more effectual or correlative than average annual fluctuations in precipitation.

If correlations between loci and physical variables are indicative of an adaptive influence (e.g., Bonin et al. 2009), they are not necessarily representative of direct causal relationships (e.g., selection for specific phenotypes). Note that from among our panel, the locus *Omy_stat3-273* was previously identified among desert and montane resident *O. mykiss* populations as being associated with temperature (Narum et al. 2010c). We demonstrated in our evaluation that this

locus was correlated with precipitation but not with temperature, yet in actuality it is likely

associated with overall climate and thus affected by myriad aspects of habitat variability. We

identified patterns of putatively adaptive variation associated with climate, but corresponding

phenotypic trait variation was not measured. However, distinguishing whether phenotypic

changes are genetically based or the result of phenotypic plasticity has proven difficult (Merila

and Hendry 2014). Often many genes are involved in adaptive responses to specific

environments and/or climates (Kassahn et al. 2007), and controlled experiments would be

necessary to make direct inferences on interactions between environments, phenotypes, and

specific genes. Rather than providing unequivocal evidence of adaptation on the basis of

phenotypic attributes, our study identifies locus associations that can be seen as indicators of

related but undetermined causative environmental forces, such as early seasonal onset (Bradshaw

and Holzapfel 2008). For example, distinct populations of steelhead trout in Oregon's Hood

River occupy either glacial fed or spring fed tributaries (Underwood et al. 2003; Matala et al

2009). Distinctions likely arose due in part to selection in variable environments, characterized

by in stream flow rate, thermal stability, and other stressors (Lytle and Poff 2004). When altered

by the forces of climate change (e.g. lengthening seasons) those environments may in turn affect

phenotypes such as spawn timing or migration timing (Crozier et al. 2011; Reed et al. 2011).

Nevertheless, it cannot be stated unequivocally that climate associations are indicative of loci

under direct selection.

Global climate change and effects of climate on ecosystems has earned the attention of

the scientific community and freshwater fish are expected to be negatively impacted (e.g.,

McCullough 2009). There is growing consensus that habitats occupied by salmonid species in

North America and Europe will, or have already experienced climate related alterations (Crozier

et al. 2006; Battin et al. 2007; Winfield et al. 2010; Neilsen et al. 2013). Most of the Columbia River Basin has been identified as habitat at high risk for thermal stress in salmonids (Wu et al. 2012), and conservation of many populations is already warranted (Busby 1996; ICTRT 2003; USOFR 2006). Climate altered habitats may lead to rapid evolution (Barrett 2010) or shifting range margins of myriad species (Hickling et al. 2005; Chen et al. 2011). In fishes this may manifest as range contractions in cold environments or expansions in warm environments, and adaptive versus neutral genetic divergence is likely to occur at differing spatial and temporal scales (Conover et al. 2006).

Therefore, conservation strategies based heavily on neutral genetic variation, with an under-emphasis on the distribution of non-neutral variation, risk detrimental impacts to locally adapted population segments, and should be scrutinized (Pearman 2001, Schwartz et al. 2009). We concur with Funk et al. (2012) that neutral and non-neutral elements of differentiation are not mutually exclusive, and should be used in concert to provide a cautious and conscientious description of species diversity in a management framework. Monitoring trends between neutral and non-neutral differentiation and the corresponding degree of disparity observed over time may be fundamentally important for addressing impacts of climate change (i.e., adaptive responses). This will conceivably have a potential role in reshaping conservation boundaries to safeguard species diversity. The design of our steelhead trout study follows this perspective; identifying selection candidate loci and environmental associations, then drawing comparisons with patterns of neutral population divergence. We observed non-neutral genetic heterogeneity of populations in association with environment. However, in our results neutral diversity encapsulated overall steelhead trout diversity with better clarity and finer resolution across established DPSs. Nevertheless, monitoring of historical and/or contemporary non-neutral

differentiation of populations adds a unique dimension to the characterization of these populations (Willing et al. 2010). The candidate markers identified in our study are expected to be useful for modeling population level responses to climate change, and future population genomics approaches with thousands of steelhead SNPs should provide improved estimation and resolution of adaptive differentiation in the Columbia River Basin.

There is a palpable consensus among researchers and managers cautioning against broad scale, rigid approaches to conservation, and acknowledging the need to address productivity limitations of myriad organisms in need of protection (Clostio et al. 2012). However, conflicting opinions on the role of genetics in conservation still pervade management agendas (Waples 1995; Fraser and Bernatchez 2001; Garcia de Leaniz et al. 2007; Allendorf et al. 2010). Maintaining overall conservation unit viability must necessarily account for the viability of all demographically important population components within those units (Crandall et al. 2000; Latch et al. 2011), particularly where demographic instability (e.g., genetic drift in small populations) may reduce overall adaptive variation or adaptive potential within regions (Kawecki & Ebert 2004; Schoville et al. 2012). It is likely that the relevance and contribution of non-neutral variation is frequently overlooked or underutilized in conservation planning, but given recent calls for the incorporation of climate science in application of the ESA (e.g., McClure et al. 2013), the utility of such information should be highlighted. In the absence of efforts to regularly evaluate putatively adaptive population differences, there is presumably a greater risk for the loss of genetic diversity as climates and habitats continue to change through time. Over the long-term, the adaptive potential of many species across taxa will need to be further explored and considered in conservation planning. The real effects of a changing environment, including

shifting ranges, may not be uniformly realized or fit tightly into predefined units (e.g., ESU,

DPS, MPG).

Acknowledgements

References

Abadía-Cardoso, A., A. J. Clemento, and J. C. Garza. 2011. Discovery and Characterization of Single-Nucleotide Polymorphisms in Steelhead/rainbow Trout, *Oncorhynchus mykiss*. Molecular Ecology Resources 11(Suppl 1):31–49.

Ackerman, M. W., W. D. Templin, J. E. Seeb, and L. W. Seeb. 2012a. Landscape heterogeneity and local adaptation define the spatial genetic structure of Pacific salmon in a pristine environment. Conservation Genetics DOI 10.1007/s10592-012-0401-7.

Ackerman, M., and M. Campbell. 2012b. Chinook and steelhead genotyping for genetic stock identification at lower granite dam. ANNUAL PROGRESS REPORT July 1, 2011 - June 30, 2012, Submitted to Bonneville Power Administration. Contract # 53239; Project # 2010-026-00. Available at: https://pisces.bpa.gov/release/documents/documentviewer.aspx?doc=P128035

Aguilar, A. and J. C. Garza. 2008. Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). Molecular Ecology Resources 8:659-662.

Akey, J.M. 2009. Constructing genomic maps of positive selection in humans: Where do we go from here? Genome Research **19**:711-722.

Allendorf, F. W., P. A. Hohenlohe and G. Luikart. 2010. Genomics and the future of conservation genetics. Nat. Rev. Genet. **11**:697-709.

Anderson, M. J., and C. J. F. ter Braak. 2003. Permutation tests for multi-factorial analysis of variance. Journal of Statistical Computation and Simulation **73**:85-113.

Anderson, M. J., R. B. Ford, D. A. Feary and C. Honeywill. 2004. Quantitative measures of sedimentation in an estuarine system and its relationship with intertidal soft-sediment infauna. Marine Ecological-Progress Series **272**:33-48.

Balkenhol, N., L. P. Waits and R. J. Dezzani. 2009. Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. Ecography **32**:818-830.

Barrett, R. D., A. Paccard, T. M. Healy, S. Bergek, P. M. Schulte, D. Schluterand S. M. Rogers. 2010. Rapid evolution of cold tolerance in stickleback. Proceedings of the Royal Society B: Biological Sciences 278:233–238.

Battin, J.; Wiley, M.W.; Ruckelshaus, M.H.; Palmer, R.N.; Korb, E.; Bartz, K.K.; Imaki, H. 2007. Projected impacts of climate change on salmon habitat restoration. Proceedings of the National Academy of Sciences 10.1073/pnas.0701685104.

Beacham, T. D., S. Pollard, and K. D. Le. 1999. Population structure and stock identification of steelhead in southern British Columbia, Washington, and the Columbia River based on microsatellite DNA variation. Transactions of the American Fisheries Society **128**:1068–1084.

Beaumont, M. A. 2005. Adaptation and speciation: what can Fst tell us? Trends in Ecology and Evolution **20**(8):435-440.

Beaumont, M.A., and D.J. Balding. 2004. Identifying adaptive genetic divergence among populations from genome scans. Molecular Ecology **13**:969-980.

Beaumont, M.A., and R.A. Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society Biological Sciences **263**:1619–1626.

Behnke, R. J. Behnke, R. J. 2002. Trout and Salmon of North America. Free Press, Simon and Shuster, Inc. N.Y., N.Y. pages 75-86.

Benjamini Y., and D. Yekutieli. 2001. The control of false discovery rate under dependency. Annals of Statistics **29**:1165-1188.

Blankenship, S. M., M. R. Campbell, J. E. Hess, M. A. Hess, T. W. Kassler, C. C. Kozfkay, A. P. Matala, S. R. Narum, M. M. Paquin, M. P. Small, J. J. Stephenson, K. I. Warheit and P. Moran. 2011. Major Lineages and Metapopulations in Columbia River Oncorhynchus mykiss Are Structured by Dynamic Landscape Features and Environments. Transactions of the American Fisheries Society **140**:665–684.

Bonin, A., F. Nicole, F. Pompanon, C. Miaud and P. Taberlet. 2007. Population adaptive index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. Conservation Biology **21**(3):697-708.

Bonin, A., P. Taberlet, C. Miaud, and F. Ponpanon. 2009. Exploratory genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria). Molecular Biology and Evolution **23**:773-783.

Bourret, V., M. Dionne, M. P. Kent, S. Lien and L. Bernatchez. 2013. Landscape genomics in Atlantic salmon (Salmo saloar): searching for gene-environment interactions driving local adaptation. Evolution doi:10.1111/evo.12139.

Bradshaw, W. E. and C. M. Holzapfel. 2008. Genetic responses to rapid climate change: it's seasonal timing that matters. Molecular Ecology 17:157–166.

Brunelli, J. P., K. J. Wertzler, K. Sundin and G. H. Thorgaard. 2008. Y-specific sequences and polymorphisms in rainbow trout and Chinook salmon. Genome 51:739-748.

Busby, P. J., T. C. Wainwright, G. J. Bryant, L. J. Lierheimer, R. S. Waples, F. W. Waknitz, and I. V. Lagomarsino. 1996. Status review of West Coast steelhead from Washington, Idaho, Oregon, and California. Technical Memorandum NMFS-NWFSC-27. Available at: www.nwr.noaa.gov.

Campbell, N. R., and S. R. Narum. 2008. Identification of novel SNPs in Chinook salmon and variation among life history types. Transactions of the American Fisheries Society **137**:96-106.

Campbell, N. R. and S. R. Narum. 2009. Identification and characterization of heat shock response-reltaed single-nucleotide polymorphisms in O. mykiss and O. tshawytscha. Molecular Ecology Resources 9:1460-1559

Campbell, N. R., K. Overturf and S. R. Narum. 2009. Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. Molecular Ecology Resources 9:318-322.

Campbell, N. R., S. J. Amish, V. L. Pritchard, K. M. MCKelvey, M. K. Young, M. K. Schwartz, J. C. Garza, G. Luikart, and S. R. Narum. 2012. Development and evaluation of 200 novel SNP assays for population genetic studies of westslope cutthroat trout and genetic identification of related taxa. Molecular Ecology Resources doi: 10.1111/j.1755-0998.2012.03161.x

Carmichael, L. E., Krizan, J., Nagy, J. A., Fuglei, E., Dumond, M., Johnson, D., Veitch, A., Berteaux, D. & Strobeck, C. (2007). Historical and ecological determinants of genetic structure in arctic candies. Molecular Ecology **16:**3466–3483.

Chen, I-Ching, J. K. Hill, R. Ohlemüller, D. B. Roy and C. D. Thomas. 2011. Rapid range shifts of species associated with high levels of climate warming. Science **333**:1024-1026.

Chilcote, M. W. 1998. Conservation status of steelhead in Oregon. Information Report 98-1.Oregon Department of Fish and Wildlife, Portland, 108p.

Chilcote, M. W., S. A. Leider, and J. J. Loch. 1986. Differential reproductive success of hatchery and wild summer-run steelhead under natural conditions. Transactions of the American Fisheries Society **115**:726–735.

Clostio, R. W., A. M. Martinez, K. E. LeBlanc and N. M. Anthony. 2012. Population genetic structure of a threatened tortoise across the south-eastern United States: implications for conservation management. Animal Conservation **15**:613-625.

Coleman, M. A. and K. D. Fausch. 2007. Cold summer temperature limits recruitment of age-0 cutthroat trout in high-elevation Colorado streams. Transactions of the American Fisheries Society **136**:1231–1244.

Conover, D.O., L. M. Clarke, S. B. Munch, and G. N. Wagner. 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. Journal of Fish Biology 69:21–47.

Crozier, L.G.; Zabel, R.W. 2006. Climate impacts at multiple scales: evidence for differential population responses in juvenile Chinook salmon. Journal of Animal Ecology **75**: 1100-1109.

Crozier, L. G., M. D. Scheuerell, and R. W. Zabel 2011. Using time series analysis to characterize evolutionary and plastic responses to environmental change: a case study of a shift toward earlier migration date in sockeye salmon. American Naturalist **178**:755–773.

Currens, K. P., C. B. Schreck and H. W. Li. 2009. Evolutionary ecology of redband trout. Transactions of the American Fisheries Society **138**:797–817.

Dionne M., F. Caron, J. J. Dodson and L. Bernatchez. 2008. Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. Molecular Ecology **17**:2382-2396.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology **14**:2611-20.

Excoffier L., G. Laval and S. Schneider. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online **1**:47-50.

Farrell, A. P., S. G. Hinch, S. J. Cooke, D. A. Patterson, G. T. Crossin, M. Lapointe and M. T. Mathes. 2008. Pacific Salmon in Hot Water: Applying Aerobic Scope Models andBiotelemetry to Predict the Success of Spawning Migrations. Physiological and Biochemical Zoology **81**(6):697–708

Felsenstein, J. 2008. PHYLIP (phylogeny inference package), version 3.68. University of Washington, Seattle.

Ford, M. J. (ed.). 2011. Status review update for Pacific salmon and steelhead listed under the Endangered Species Act: Pacific Northwest. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-113, 281 p.

Fowler S. L., D. Hamilton and S. Currie. 2009. A comparison of the heat shock response in juvenile and adult rainbow trout (*Oncorhynchus mykiss*) – implications for increased thermal sensitivity with age. Canadian Journal of Fisheries and Aquatic Sciences **66**:91-100.

Fraser, D. J. and L. Bernatchez. 2001. Adaptive evolutionary conservation: towards a unified concept for defining conservation units. Molecular Ecology **10**:2741-2752.

Fraser, D. J., L. K. Weir, L. Bernatchez, M. M. Hansen and E. B. Taylor. 2011. Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. Heredity **106**:404-420.

Fryer, J. K., J. Whiteaker and D. Kelsey. 2012. Upstream Migration Timing of Columbia Basin Chinook Salmon, Sockeye Salmon, and Steelhead in 2010. Columbia River Inter-Tribal Fish Commission Technical Report for BPA Project 2008-503-00. March 31, 2012.

Funk, W. C., J. K. McKay, P. A. Hohenlohe and F. W. Allendorf. 2012. Harnessing genomics for delineating conservation units. Trends in Ecology and Evolution **27**(9):489-496.

Garcia de Leaniz, C., I. A. Fleming, S. Einum, E. Verspoor, W. C. Jordan, S. Consuegra, N. Aubin-Horth, D. Lajus, B. H. Letcher, A. F. Youngson, J. H. Webb, L. A. Vøllestad, B. Villanueva, A. Ferguson and T. P. Quinn. 2007. A critical review of adaptive genetic variationin Atlantic salmon: implications for conservation. Biological Reviews **82**:173-211.

Ge, Song, De-yuan Hong, Hai-qun Wang, Zheng-yu Liu and Can-ming Zhang. 1998. Population genetic structure and conservation of an endangered conifer, Cathaya argyrophylla (Pinaceae). International Journal of Plant Sciences **159**(2):351-357.

Good, T.P., R.S. Waples, and P. Adams (editors). 2005. Updated status of federally listed ESUs of West Coast salmon and steelhead. U.S. Dept. Commerce, NOAA Tech. Memo. NMFS-NWFSC-66, 598 p.

Goode, J. R., J. M. Buffington, D. Tonina, D. J. Isaak, R. F. Thurow, S. Wenger, D. Nagel, C. Luce, D. Tetzlaff and C. Soulsby. 2013. Potential effects of climate change on streambed scour and risks to salmonid survival in snow-dominated mountain basins. Hydrological Processes **27**:750–765.

Hansen, M. H., S. Young, H. B. H. Jorgensen, C. Pascal, M. Henryon, and J. Seeb. 2011. Assembling a dual purpose TaqMan-based panel of single-nucleotide polymorphism markers in

rainbow trout and steelhead (*Oncorhynchus mykiss*) for association mapping and population genetics analysis. Molecular Ecology Resources 11(Suppl. 1):67–70.

Hari, R. E., D. M. Livingstone, R. Siber, P. Burkhardt-Holm and H. Guttinger. 2006. Consequences of climatic change for water temperature and brown trout populations in Alpine rivers and streams. Global Change Biology **12**:10-26.

Hendry, M. A., J. K. Wenburg, K. W. Myers, and A. P. Hendry. 2002. Genetic and phenotypic variation through the migratory season provides evidence for multiple populations of wild steelhead in the Dean River, British Columbia. Transactions of the American Fisheries Society **131**:418–434.

Hermisson, J. 2009. Who believes in whole-genome scans for selection? Heredity **103**:283-284.

Hess, J. E., N. R. Campbell, A. P. Matala and S. R. Narum. 2011. Genetic Assessment of Columbia River Stocks. Columbia River Inter-Tribal Fish Commission annual report 2010. Project Number: 2008-907-00, Contract Number: 41224, Submitted to Bonneville Power Administration, February 16, 2011.

Hess, J. E., N. R. Campbell, A. P. Matala and S. R. Narum. 2012. Genetic Assessment of Columbia River Stocks. Columbia River Inter-Tribal Fish Commission annual report 2011. Project Number: 2008-907-00, Contract Number: 41224, Submitted to Bonneville Power Administration, March 31, 2012.

Hickling, R., D. B. Roy, J. K. Hill and C. D. Thomas. 2005. A northward shift of range margins in British Odonata. Global Change Biology **11**:502-506.

Hoffmann, A.A. and C. M. Sgro`. 2011. Climate change and evolutionary adaptation. Nature 470:479–485.

Holderegger, R., and H. H. Wagner. 2008. Landscape Genetics. Bioscience **58**(3):199-207.

ICTRT (Interior Columbia Technical Recovery Team). 2003. Independent Populations of Chinook, steelhead and sockeye for listed Interior Columbia Basin ESUs. Interior Columbia Basin Technical Recovery Team Report. July 2003.

Isaak, D. J., C. C. Muhlfeld, A. S. Todd, R. Al-Chokhachy, J. Roberts, J. L. Kershner, K. D. Fausch and S. W. Hostetler. 2012a. The past as prelude to the future for understanding 21st-century climate effects on Rocky Mountain trout. Fisheries **37**(12):542-556.

Isaak, D. J., S. Wollrab, D. Horan and G. Chandler. 2012b. Climate change effects on stream and river temperatures across the northwest U.S. from 1980-2009 and implications for salmonid fishes. Climate Change 113(2):499-524.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics **23**:1801–1806.

Kassahn K. S,, R. H. Crozier, A. C. Ward, G. Stone and M. J. Caley. 2007. From transcriptome to biological function: environmental stress in an ectothermic vertebrate, the coral reef fish *Pomacentrus moluccensis*. BMC Genomics, doi:10.1186/1471-2164-8-358

Kawecki, T. J. and D. Ebert. 2004. Conceptual issues in local adaptation. Ecology Letters **7**:1225–1241.

Kovach, R. P., A. J. Gharrett and D. A. Tallmon. 2012. Genetic change for earlier migration timing in a pink salmon population. Proceedings of the Royal Society **279**:3870–3878

LaHood, E.S., J.J. Miller, C. Apland, M.J. Ford. 2008. A rapid, ethanol-free fish tissue collection method for molecular genetic analyses. Transactions of the American Fisheries Society **137**:1104-1107.

Landguth, E.L. and N. Balkenhol. 2012. Relative sensitivity of neutral versus adaptive genetic data for assessing population differentiation. Conservation Genetics **13**(5):1-6.

Latch, E. K., W. I. Boarman, A. Walde and R. C. Fleischer. 2011. Fine-scale analysis reveals cryptic landscape genetic structure in desert tortoises. PLos ONE **6**(11):e27794.

Limborg, M. T., S. M. Blankenship, S. F. Young, F. M. Utter, L. W. Seeb, M. H. H. Hansen, and J. E. Seeb. 2011. Signatures of natural selection among lineages and habitats in *Oncorhynchus mykiss*. Ecology and Evolution. doi:10.1002/(ISSN)2045-7758

Lin, Long-Hui, Yan-Fu Qu, Hong Li, Kai-Ya Zhou and Xiang Ji. 2012. Genetic structure and demographic history should inform conservation: Chinese cobras currently treated as homogenous show population divergence. PLos ONE **7**(4): e36334. doi:10.1371/journal.pone.0036334

Loarie, S.R., P. B. Duffy, H. Hamilton, G. P. Asner, C. B. Field and D. D. Ackerly. 2009. The velocity of climate change. Nature 462:1052–1055.

Lytle, D. A. and N. L. Poff. 2004. Adaptation to natural flow regimes. Trends in ecology and evolution **19**(2):94-100.

Manel, S., M. K. Schwartz, G. Luikart and P. Taberlet. 2003. Landscape genetics: combining landscape ecology and population genetics. Trends in Ecology and Evolution **18**(4):189-197.

Matala, A. P., S.Marx, and T. G.Wise. 2008. A genetically distinct wild redband trout (*Oncorhynchus mykiss gairdneri*) population in Crane Prairie Reservoir, Oregon, persists despite

extensive stocking of hatchery rainbow trout (*O. m. irideus*). Conservation Genetics **9**:1643–1652.

Matala, A. P., R. French, E. Olsen, and W. R. Ardren. 2009. Ecotype distinctions among steelhead in Hood River, Oregon, allow real-time genetic assignment of conservation broodstocks. Transactions of the American Fisheries Society **138**:1490–1509.

Matala, A. P., J. E. Hess and S. R. Narum. 2011. Resolving adaptive and demographic divergence among Chinook salmon populations in the Columbia River basin. Transactions of the American Fisheries Society **140**:783–807.

McArdle, B. H. and M. J. Anderson. 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology **82**(1):290-297.

McClure, M. M., M. Alexander, D. Borggaard, D. Boughton, L. Crozier, R. Griffis, J. C. Jorgensen, S. T. Lindley, J. Nye, M. J. Rowland, E. E. Seney, A. Snover, C. Toole and K. VanHoutan. 2013. Incorporating Climate Science in Applications of the U.S. Endangered Species Act for Aquatic Species. Conservation Biology 27(6):1222–1233.

McCullough, D. A., J. M. Bartholow, H. I. Jager, R. L. Beschta, E. F. Cheslak, M. L. Deas, J. L. Ebersole, J. S. Foott, S. L. Johnson, K. R. Marine, M. G. Mesa, J. H. Petersen, Y. Souchon, K. F. Tiffan and W. A. Wurtsbaugh. 2009. Research in Thermal Biology: Burning Questions for Coldwater Stream Fishes. Reviews in Fisheries Science **17**(1):90-115.

Merila, J., and A. Hendry. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evolutionary Applications 7:1–14.

Milner, A., A. L. Robertson, K. A. Monaghan, A. J. Veal and E. A. Flory. 2008. Colonization and development of an Alaskan stream community over 28 years. The Ecological Society of America **6**(8):413-419.

Narum, S. R., J. Zendt, D. Graves, and B. Sharp. 2008. Influence of landscape on resident and anadromous life history types of *Oncorhynchus mykiss*. Canadian Journal of Fisheries and Aquatic Sciences **65**:1013-1023.

Narum, S. R., N. Campbell, A. P. Matala and J. E. Hess. 2010a. Genetic assessment of Columbia River Stocks: 2009 annual report. Submitted to Boneville Power Administration. Contract #41224; Project # 2008-907-00. Available at: https://pisces.bpa.gov/release/documents/documentviewer.aspx?doc=P115281

Narum, S. R., N. R. Campbell, C. C. Kozfkay and K. A. Meyer. 2010b. Adaptation of redband trout in desert and montane environments. Molecular Ecology doi: 10.1111/j.1365-294X.2010.04839.x.

Narum, S. R., J. E. Hess and A. P. Matala. 2010c. Examining Genetic Lineages of Chinook Salmon in the Columbia River Basin. Transactions of the American Fisheries Society DOI: 10.1577/T09-150.1.

Narum, S. R., J. S. Zendt, C. Frederiksen, N. Campbell, A. Matala, and W. Sharp. 2011. Candidate Genetic Markers Associated with Anadromy in *Oncorhynchus mykiss* of the Klickitat River. Transactions of the American Fisheries Society **140**(3):843-854.

Narum, S. R. and J. Hess. 2011. Comparison of FST outlier tests for SNP loci under selection. Molecular Ecology Resources **11**(Suppl. 1):184–194.

Narum, S. R., N. R. Campbell, K. A. Meyer, M. R. Miller and R. W. Hardy. 2013. Thermal adaptation and acclimation of ectotherms from differing aquatic climates. Molecular Ecology **22**:3090–3097.

Nei, M. 1972. Genetic distance between populations. The American Naturalist **106**:283–292.

Nielsen, J. L., A. Byrne, S. L. Graziano and C. C. Kozfkay. 2009. Steelhead genetic diversity at multiple spatial scales in a managed basin: Snake River, Idaho. North American Journal of Fisheries Management **29**:680-701.

Nielsen, J. L., G. T. Ruggerone, and C.E. Zimmerman. 2013. Adaptive strategies and life cycle characteristics in a warming climate: salmon in the Arctic? Environmental Biology of Fishes. 96(10-11):1187-1226

Nosil, P., S. P. Egan and D. J. Funk. 2007. Heterogeneous genomic differentiation between walking-stick ecotypes: "isolation by adaptation" and multiple roles for divergent selection. Evolution 62-2:316–336.

Olsen, J. B., T. D. Beacham, M Wetklo, L. W. Seeb, C. T. Smith, B. G. Flannery and J. K. Wenburg. 2010. The influence of hydrology and waterway distance on population structure of Chinook salmon *Oncorhynchus tshawytscha* in a large river. Journal of Fish Biology **76**:1128-1148.

Olsen, J. B., P. A. Crane, B. G. Flannery, K. Dunmall, W. D. Templin and J. K. Wenburg. 2011. Comparative landscape genetic analysis of three Pacific salmon species from subarctic North America. Conservation Genetics **12**:223–241.

Page, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences **12**: 357-358.

Peakall, R., and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6, 288-295.

Program note available from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1471-8286.2005.01155.x

Pearman, P. B. 2001. Conservation value of independently evolving units: sacred cow or testable hypothesis? Conservation Biology **15**(3):780-783.

Perry, G. M. L., R. G. Danzmann, M. M. Ferguson, and J. P. Gibson. Quantitative trait loci for upper thermal tolerance in outbred strains of rainbow trout (*Oncorhynchus mykiss*). Heredity **86**:333–341.

Pritchard, J., M. Stephens, P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics **155**:945-959.

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

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution **43**:223-225.

Reed, T. E., D. Schindler, M. Hague, D. Patterson, E. Meir, R. S. Waples, and S. Hinch. 2011. Time to evolve? potential evolutionary responses of Fraser River sockeye salmon to climate change and effects on persistence. PLoS ONE 6:e20380.

Rieman, B. E., D. Isaak, S. Adams, D. Horan, D. Nagel, C. Luce and D. Myers. 2007. Anticipated climate warming effects on bull trout habitats and populations across the interior Columbia River Basin. Transactions of the American Fisheries Society **136**:1552-1565.

Ruokonen, M., L. Kvist, T. Aarvak, J. Markkola, V. V. Morozov, I. J. Oien, E. E. Syroechkovsky Jr., P. Tolvanen and J. Lumme. 2004. Population genetic structure and conservation of the lesser white-fronted goose *Anser erythropus*. Conservation Genetics 5:501-512.

Sanchez, C. C., T. P. L. Smith, R. T. Wiedmann, R. L. Vallejo, M. Salem, J. Yao and C. E. Rexroad III. 2009. Single nucleotide polymorphism discovery in rainbow trout by deep sequencing of a reduced representation library. BMC Genomics 10:559.

Schwartz, M. K., G. Luikart, K. S. McKelvey, and S. A. Cushman. 2009. Landscape genomics: a brief perspective. In: Cushman, S. A., Huettmann F. (eds) Spatial Complexity, informatics, and wildlife conservation. Springer, Berlin.

Schoville, S. D., A. Bonin, O. Francois, S. Lobreaux, C. Melodelima and S. Manel. 2012. Adaptive genetic variation on the landscape: methods and cases. Annual Reviews of Ecology Evolution and Systematics **43**:23-43.

Scott, J.B., Jr., and Gill, W.T. (*Editors*). 2008. *Oncorhynchus mykiss*: assessment of Washington State's anadromous populations and programs. Washington Department of Fish and Wildlife,Olympia, Wash.

Segelbacher, G., S. A. Cushman, B. K. Epperson, MJ. Fortin, O. Francois, O. J. Hardy, R. Holderegger, P. Taberlet, L. P. Waits and S. Manel. 2010. Applications of landscape genetics in conservation biology: concepts and challenges. Conservation Genetics **11**:375-385.

Sepulveda-Villet, O. J. and C. A. Stepien. 2012. Waterscape genetics of the yellow perch (Perca flavescens): patterns across large connected ecosystems and isolated relict populations. Molecular Ecology (23):5795-5826.

Sprowles, A. E., M. R. Stephens, N. W. Clipperton and B. P. May. 2006. Fishing for SNPs: A targeted locus approach for single nucleotide polymorphism discovery in rainbow trout. Transactions of the American Fisheries Society 135:1698-1721.

Underwood, K. D., N. K. Ackerman, C. G. Chapman, K. L. Witty, S. P. Cramer, and M. L. Hughes. 2003. Hood River Production Program Review 1991–2001. Prepared by S. P. Cramer and Associates for the U.S. Department of Energy, Bonneville Power Administration, Portland, Oregon.

USOFR (U.S. Office of the Federal Register). 2006. Endangered and threatened species: final listing determinations for 10 distinct population segments of West Coast steelhead. Federal Register **71**:3(5 January 2006):834–861.

Wagner, R. S., M. P. Miller, C. M. Crisafulli and S. M. Halg. 2005. Geographic variation, genetic structure, and conservation unit designation in the Larch Mountain salamander (*Plethodon larselli*). Canadian Journal of Zoology **83**:396-406.

Waples, R. S. 1995. Evolutionarily significant units and the conservation of biological diversity under the Endangered Species Act. American Fisheries Society Symposium **17**:8-27.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution **38**:1358-1370.

Wenger, S. J., D. J. Isaak, C. H. Luce, H. M. Neville, K. D. Fausch, J. B. Dunham, D. C. Dauwalter, M. K. Young, M. M. Elsner, B. E. Rieman, A. F. Hamlet and J. E. Williams. 2011. Flow regime, temperature, and biotic interactions drive differential declines of trout species under climate change. Proceedings of the National Academy of Sciences **108**:14175–14180

Willing, E-M., P. Bentzen, C. Van Oosterhout, M. Hoffmann, J. Cable, F. Breden, D. Weigel and C. Dreyer. 2010. Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. Molecular Ecology **19**:968-984.

Winans GA, Paquin MM, Van Doornik DM (2004) Genetic stock identification of steelhead in the Columbia River basin: An evaluation of different molecular markers. North American Journal of Fisheries Management **24:**672-685.

Winfield, I. J., J. Hateley, J. M. Fletcher, J. B. James, C. W. Bean and P. Clabburn. Population trends of Arctic charr (Salvelinus alpinus) in the UK: assessing the evidence for a widespread decline in response to climate change. Hydrobiologia **650**:55-65.

Wu, H., J.S. Kimball, M.M. Elsner, N. Mantua, R.F. Adler, and J. Stanford. 2012. Projected climate change impacts on the hydrology and temperature of Pacific Northwest rivers. Water Reources Research 48:W11530, doi:10.1029/2012WR012082.

Zimmerman, C. E., and G. H. Reeves. 2000. Population structure of sympatric anadromous and nonanadromous *Oncorhynchus mykiss*: evidence from spawning surveys and otolith microchemistry. Canadian Journal of Fisheries and Aquatic Sciences **57**:2152–2162.

Zydlewski, G. B., A. Haro, and S. D. McCormick. Evidence for cumulative temperature as an initiating and terminating factor in downstream migratory behavior of Atlantic salmon (*Salmo salar*) smolts. Canadian Journal of Fisheries and Aquatic Sciences **62**(1):68–78.

Table 1. Summary of *O. mykiss* populations, sample size, and characteristics by distinct population segment (DPS) and MPG in the Columbia River Basin. All listed DPSs have an ESA status listing of threatened with the exception of the Quinault River. The Southwest Washington DPS also includes the Quinault River populations from the coast of Washington State.

DPS	tributary/ region	MPG	# pops	lineage	run	min	max	mean	total
						min	max	mean	total
Southwest WA	Columbia Estuary	n/a	4	coastal	winter	43	164	85.8	343
Lower Columbia	Clackamas	Cascade	4	coastal	mixed	43	92	60.3	241
Lower Columbia	Other	Cascade; Gorge	9	coastal	mixed	28	94	68.1	613
Upper Willamette	Willamette	Willamette	3	coastal	mixed	39	93	60.7	182
Upper Willamette	Willamette/ West Side	Willamette	3	coastal	winter	25	30	27.0	81
Middle Columbia	Klickitat	Cascade East Slope	10	inland	summer	33	48	43.4	434
Middle Columbia	Deschutes	Cascade East Slope	5	inland	summer	31	63	51.4	257
Middle Columbia	John Day	John Day	10	inland	summer	18	107	36.3	363
Middle Columbia	Yakima	Yakima	7	inland	summer	21	59	36.9	258
Middle Columbia	other	mixed	7	inland	mixed	34	148	100.1	701
Upper Columbia	Wenatchee	Upper Columbia / East Slope Cascades	6	inland	summer	19	99	40.8	245
Upper Columbia	other	Upper Columbia / East Slope Cascades	5	inland	summer	90	99	94.3	475
Snake	lower Snake	lower Snake	6	inland	summer	49	105	83.5	501
Snake	Lower Clearwater	Clearwater	5	inland	summer	49	156	107.8	539
Snake	M. F. Clearwater	Clearwater	14	inland	summer	35	99	56.4	789
Snake	S. F. Clearwater	Clearwater	4	inland	summer	36	104	57.8	231
Snake	Grande Ronde	Grande Ronde	7	inland	summer	45	95	62.7	439
Snake	Imnaha	Imnaha	4	inland	summer	23	61	41.5	166
Snake	S. F. Salmon	Salmon	4	inland	summer	39	45	43.5	174
Snake	M. F. Salmon	Salmon	8	inland	summer	23	84	48.3	386
Snake	Upper Salmon	Salmon	7	inland	summer	37	117	83.1	582
Snake	Other Salmon	Salmon	7	inland	summer	43	99	55.1	386
Snake	Hatchery	mixed	6	inland	summer	89	146	104.2	625
Total			145						9,011

Note: The "genotyped (n)" header spans the min, max, mean columns.

Table 2. Candidate SNP markers associated with climate for *O. mykiss* in the Columbia River Basin. Climate associations are: P – precipitation, T – temperature, E – elevation, D – distance. Candidate loci that were most highly correlated with temperature and/or precipitation are identified with an asterisk; see Figure 3). Notes include: inland "O_i", and coastal outlier loci "O_c", and reference numbered "precedence" loci: 1. Thermal stress association (Narum et al. 2013); 2. Temperature or precipitation association (Narum et al. 2010b). Light gray cells denote associations with a 0<BF>1, and dark gray denotes a BF>1 (P<0.01).

| SNP locus | notes | inland | | | | coastal | |
		lineage	climate associations			lineage	climate associations
Omy_hsc715-80		Candidate	P			Neutral	
Omy_SECC22b-88		Candidate	P			Neutral	
OMS00014		Candidate	P			Neutral	
OMS00062		Candidate	P			Neutral	
OMS00151		Candidate	T	P		Neutral	
Omy_97660-230		Candidate	T	D		Neutral	
Omy_CRBF1-1		Candidate	P	D		Neutral	
Omy_e1-147		Candidate	D	T		Neutral	
Omy_GHSR-121		Candidate	P			Neutral	
Omy_IL6-320		Candidate	T	P		Neutral	
Omy_metA-161		Candidate	T	P		Neutral	
Omy_nkef-241		Candidate	D			Neutral	
Omy_ntl-27		Candidate	D	P		Neutral	
Omy_u09-53.469	O_i	Candidate	T	P	D	Neutral	
Omy_UT16_2-173	O_i	Candidate	D			Neutral	
OMY1011SNP		Candidate	P			Neutral	
Omy_hsp47-86	(1)	Candidate	T	D		Ambiguous	

Locus		Classification	Associations	Classification	Associations
Omy_tlr5-205	(2)	Candidate	T P D	Ambiguous	
Omy_gdh-271	(2)	Candidate	E	Ambiguous	
Omy_97954-618	O_c	Candidate	P D	Ambiguous	
Omy_OmyP9-180	(1)	Candidate	T P	Candidate	T
Omy_stat3-273	(2)	Candidate	P	Candidate	P
Omy_aldB-165	(2)	Ambiguous		Candidate	T
Omy_hsf2-146	(2)	Ambiguous		Candidate	P D
OMS00008		Neutral		Candidate	P
OMS00058		Neutral		Candidate	T D
OMS00111		Neutral		Candidate	P D
Omy_bcAKala-380rd	O_c	Neutral		Candidate	T D
Omy_cox1-221		Neutral		Candidate	D T P

1. Thermal stress association (Narum et al. 2013); 2. Temperature or precipitation association (Narum et al. 2010b).

Figure 1. Map of the Columbia River Basin identifying locations of steelhead populations. Shading of the map reflects mean annual temperature maximums across the region from the period 1971-2000 (http://www.prism.oregonstate.edu/docs/meta/tmax_30s_meta.htm#7). Populations by DPS are: triangles – Snake River, circles – upper Columbia River, squares – middle Columbia River, stars – Upper Willamette River, and diamonds – lower Columbia River. Landmarks are: a – Quinault River, b – confluence of Willamette and Columbia rivers, c – Big White Salmon River, d – Klickitat River, e – Yakima River, f – confluence of Snake and Columbia rivers, g – confluence of Snake and Clearwater rivers, and h – confluence of Snake and Salmon rivers.
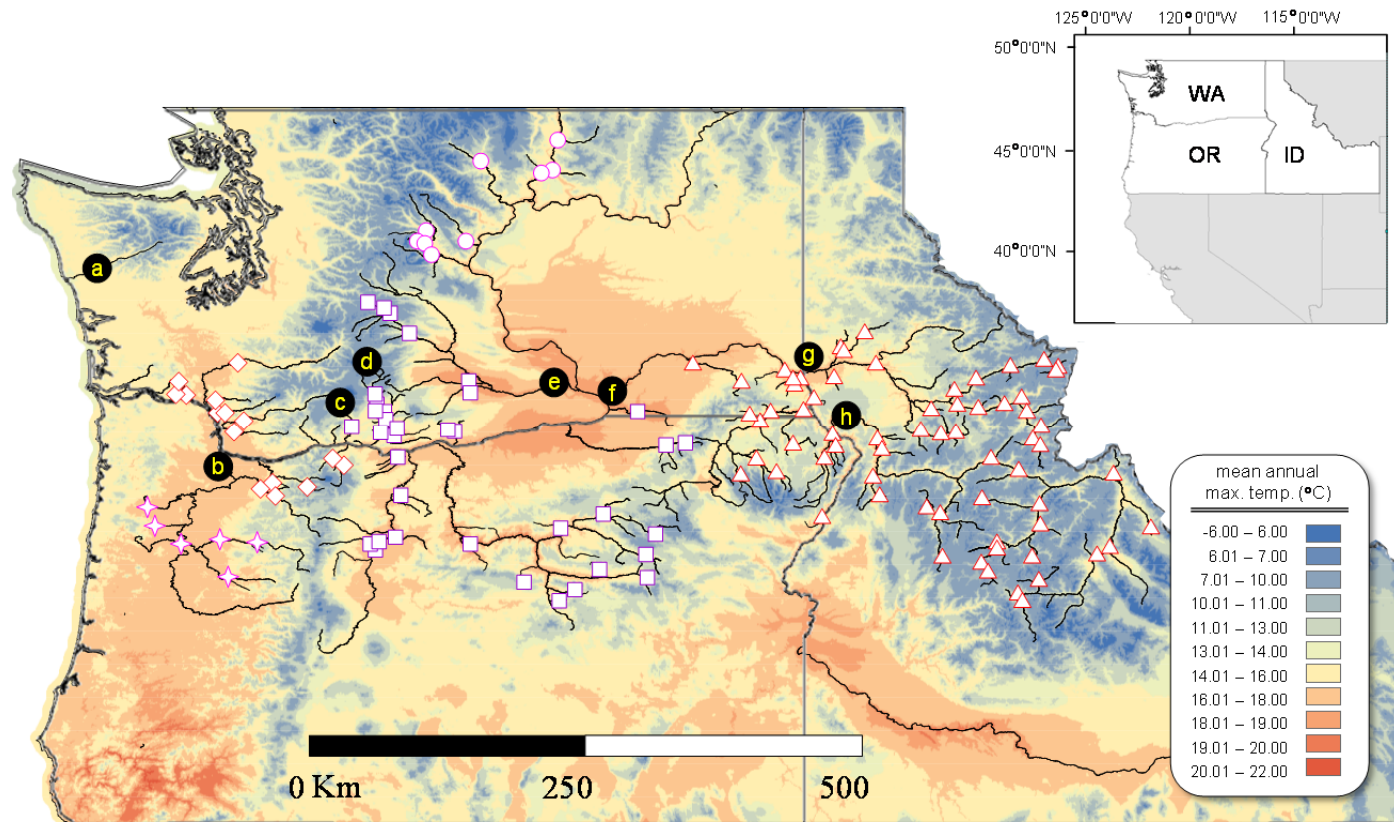
Figure 2. Mean admixture proportion plot of STRUCTURE version 2.3.4 (Pritchard et al. 2000) generated from 180 SNP loci. The histogram shows admixture proportions for *k*=2 inferred groups represented by black bars (coastal proportion) and gray bars (inland proportion). Populations correspond to reference numbers (Supplemental 1) in ascending order from left to right (except Ref.#1 Quinault in the 12th position), and are arranged by DPS. Populations at the western extreme of the inland range, and eastern extreme of the coastal range are identified in relation to (adjacent) the Cascade Mountain crest.
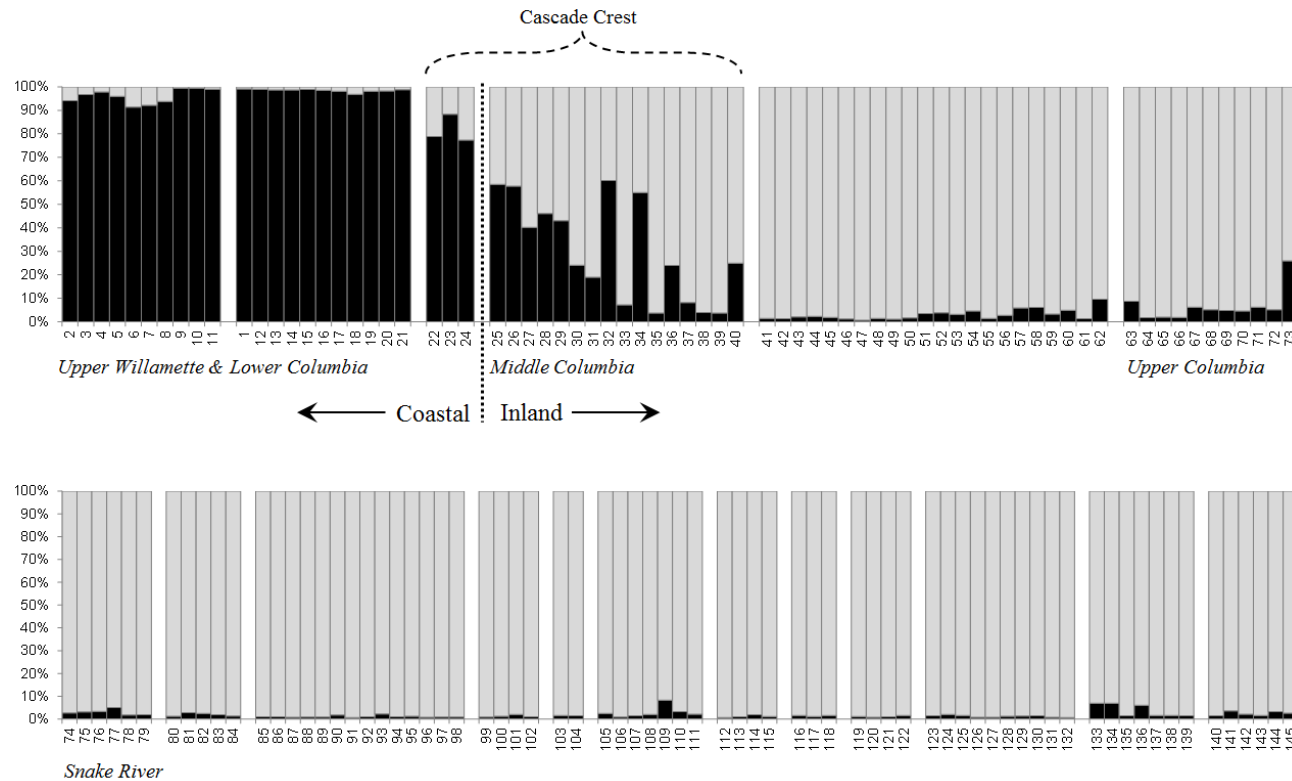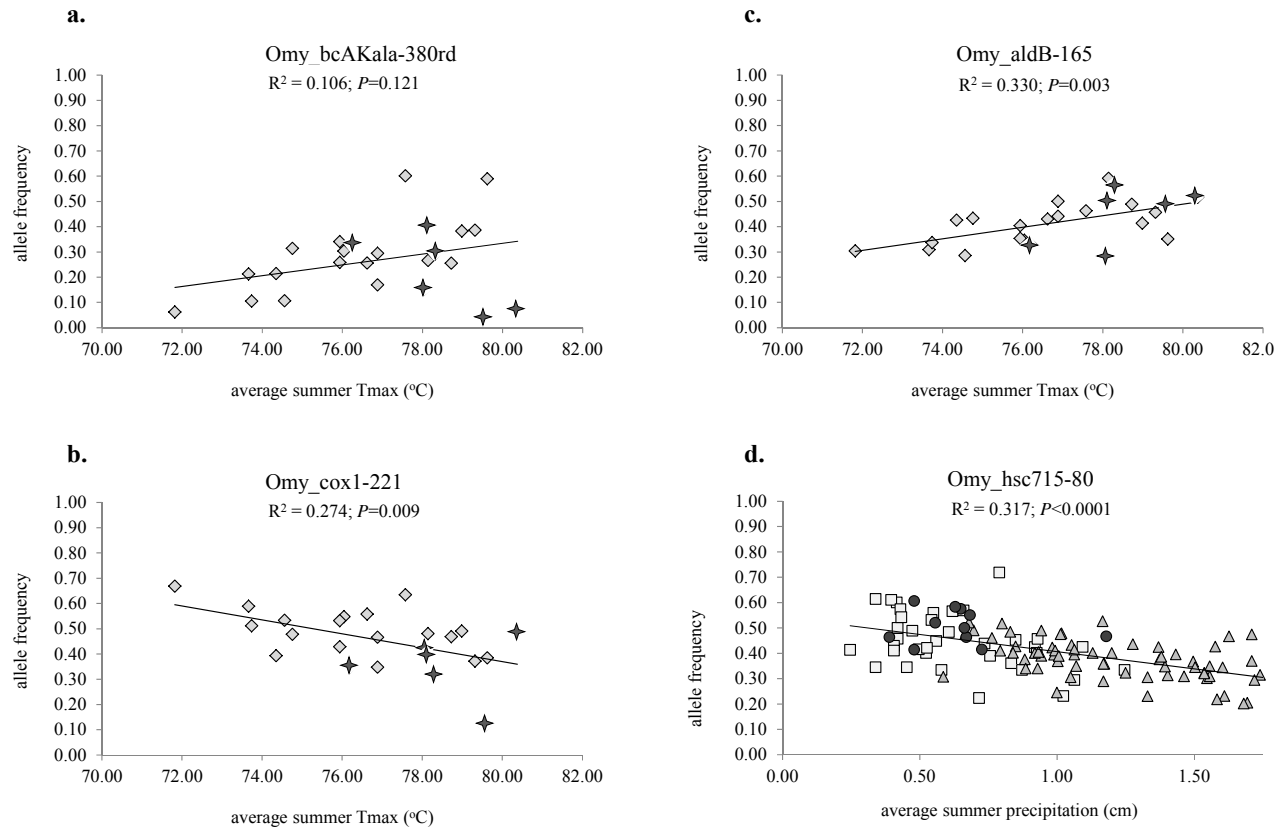
Figure 3. Linear regression plots, representing results for eight of the most highly correlated SNP loci (see Table 2). Plots a-c are temperature associations in the coastal lineage, while d-h are precipitation associations in the inland lineage. Symbols correspond with DPS: diamond – lower Columbia, star – upper Willamette, square – middle Columbia, circle – upper Columbia, and triangle – Snake.

a.

Omy_bcAKala-380rd

$R^2 = 0.106; P=0.121$

c.

Omy_aldB-165

$R^2 = 0.330; P=0.003$

b.

Omy_cox1-221

$R^2 = 0.274; P=0.009$

d.

Omy_hsc715-80

$R^2 = 0.317; P<0.0001$

e.

Omy_1011SNP

$R^2 = 0.261$; *P*<0.0001

allele frequency

average summer precipitation (cm)

g.

Omy_ntl-27

$R^2 = 0.212$; *P*<0.0001

allele frequency

average summer precipitation (cm)

f.

Omy_u09-53.469

$R^2 = 0.252$; *P*<0.0001

allele frequency

average summer precipitation (cm)

h.

Omy_97954-618

$R^2 = 0.199$; *P*<0.0001

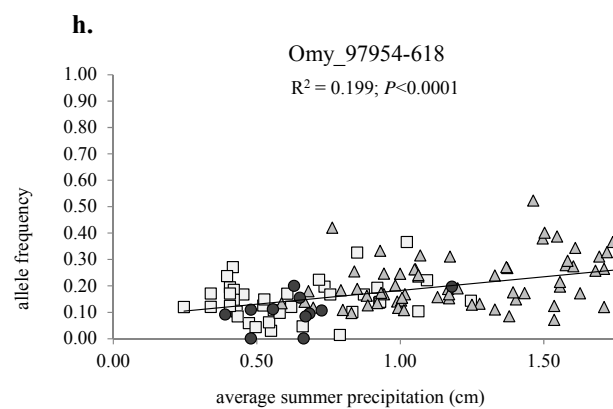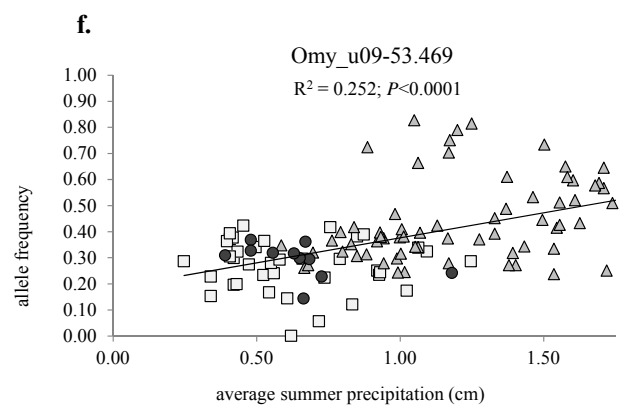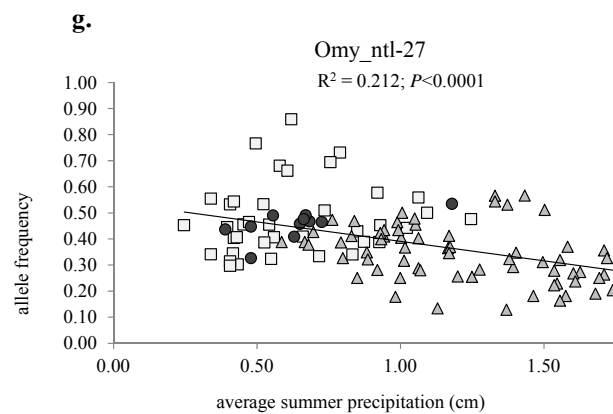allele frequency
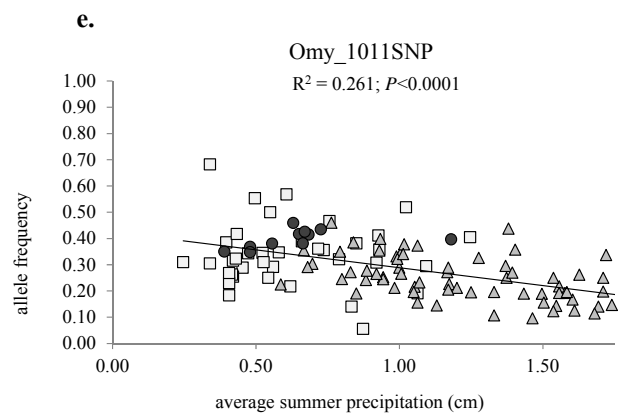
average summer precipitation (cm)

Figure 4. Neighbor joining trees depicting Nei's genetic distances among coastal lineage populations based on a) neutral variation – 158 SNPs, and b) non-neutral variation – nine SNPs. Bootstrap support exceeding 50% appears at nodes. Diamonds represent the Lower Columbia DPS, while stars represent the Upper Willamette DPS. Numbers in parentheses correspond to population reference numbers (Supplemental 1).
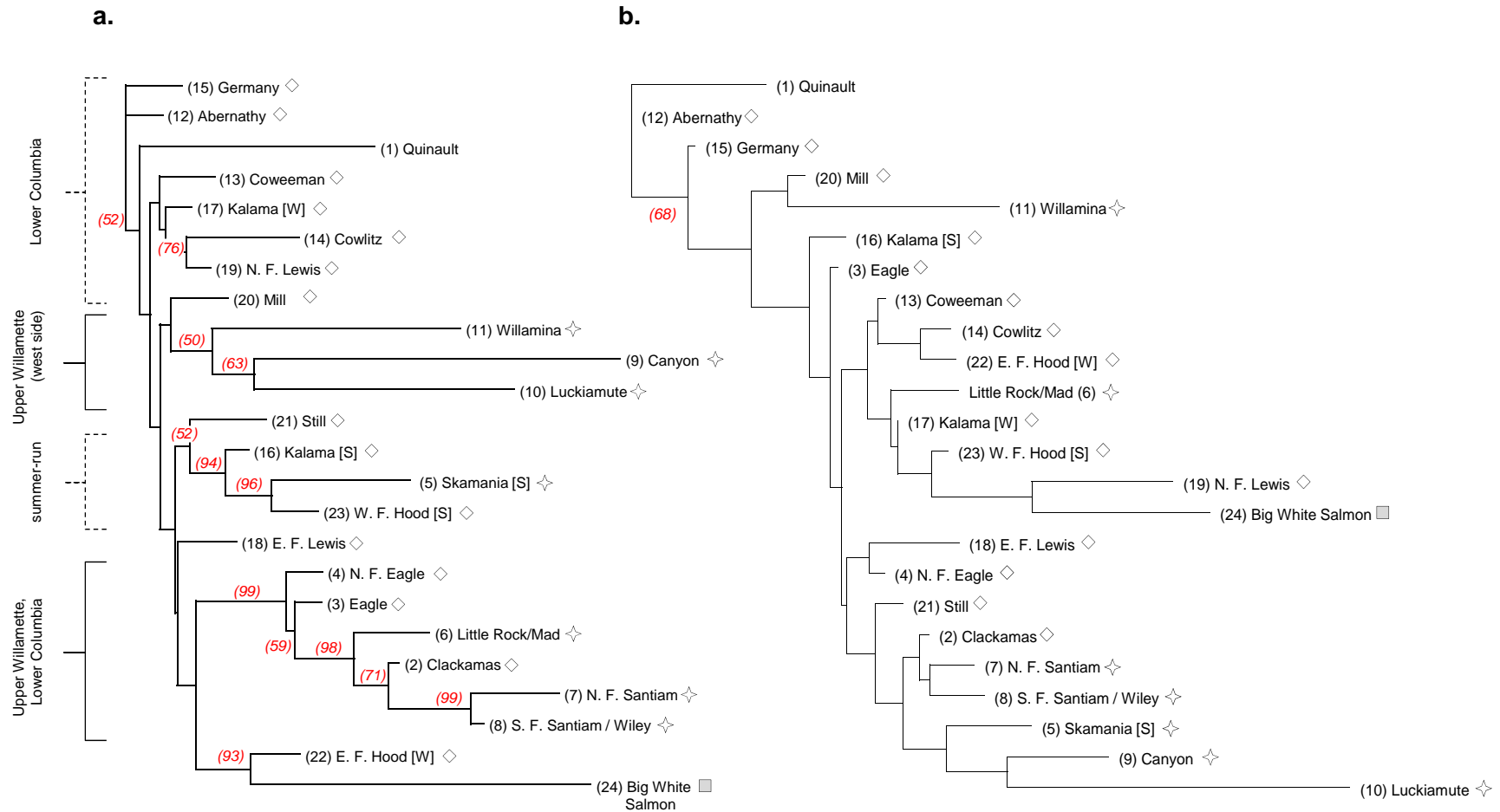
Figure 5. Neighbor joining trees depicting Nei's genetic distances among inland lineage populations. Numbers at branch ends correspond to population reference numbers (Supplemental 1) Trees are based on a) neutral variation – 146 SNPs, and b) non-neutral variation – 22 SNPs. Bootstrap support exceeding 50% appears at nodes. Population symbols correspond with DPS: triangles – Snake River, circles – upper Columbia River, squares – middle Columbia River. Major drainage sub-basins are labeled on the left. The symbol (*) represents climate related distinction among Lower Clearwater River populations and (£) represents differences among Yakima River populations based on climate. The symbol (®) identifies the branch adjoining Clearwater River and M. F. Salmon River populations. The color scale indicates population ranking by climate (1=coldest, 115=warmest) for 115 inland populations identified by reference number (Supplemental 1).
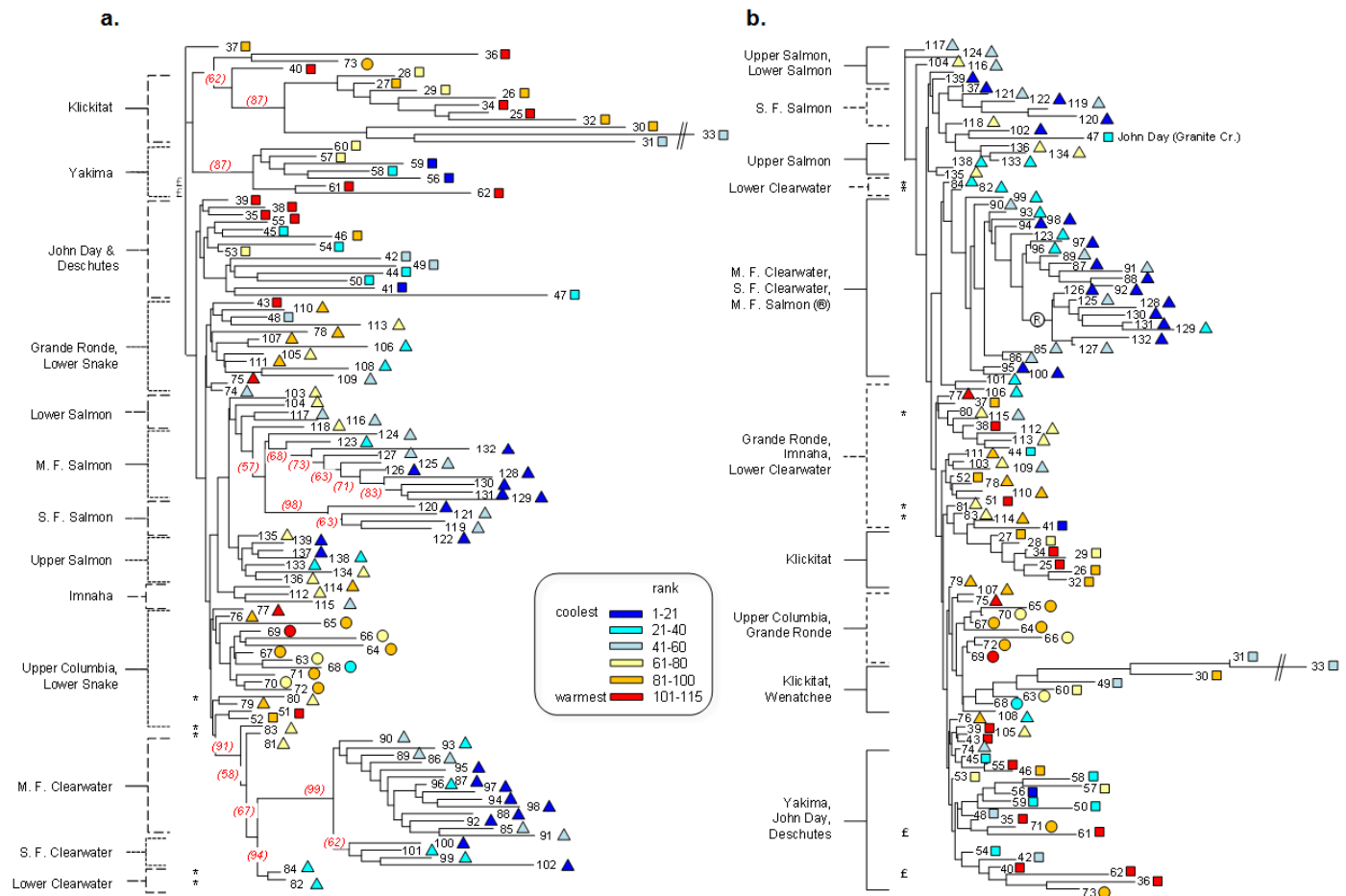
SECTION 3: Ongoing/Future Studies

As an ongoing Accords project, preparations are underway for further evaluation of multiple traits such as age-at-maturity (jacking) in Chinook salmon, run-timing in steelhead and Chinook salmon, and thermal adaptation in *O. mykiss*. More details are provided below regarding plans to investigate each of these traits in the upcoming performance period (July 1, 2014 – June 30, 2015).

Heritability and Genomic Basis for Age-at-maturity

To investigate age-at-maturity (jacking) in Chinook salmon, we are utilizing known pedigrees to evaluate heritability of jacking and the genomic basis for age-at-maturity. Migratory Chinook salmon typically reach maturity between the ages of three to five years. The three year old age class is almost exclusively comprised of males called "Jacks", which exhibit an alternative reproductive strategy. These males are smaller and less competitive breeders than older males and generally spawn with older females in a "sneaker" reproductive tactic. Overall the "Jack" life history comprises a greater proportion of the returning males in hatchery derived Chinook salmon than natural Chinook salmon in Johnson Creek, ID. The propensity to "Jack" is strongly associated with juvenile growth and rearing environment, where those males that enjoy a productive environment with plentiful resource will grow faster than they might in the wild leading to an increased likelihood to "Jack". For this study we used multi-generational pedigrees to estimate the heritability of "Jacking" in both hatchery and wild origin Johnson Creek male Chinook salmon, to determine the extent of the genetic contribution to this life history tactic.

Additionally, the known pedigree population of Chinook salmon from Johnson Creek provide the opportunity to investigate the genomic basis for age-at-maturity. Age at reproductive maturity in Chinook salmon is closely linked to overall body size at maturity and reproductive success, wherein older and larger returning Chinook salmon are capable of outcompeting younger and smaller conspecifics for reproductive opportunities. Additionally, larger fish produce more gametes, and thus are capable of producing more offspring. It is known that age at maturity is in part determined by a heritable genetic component in addition to environmental effects, and age at maturity is therefore susceptible to natural and artificial selection pressures. Many Chinook salmon hatchery programs report a shift in the age structure of their hatchery returns, with a larger proportion of younger and smaller fish returning, when compared to wild populations. It is possible that hatchery selection regimes are unintentionally selecting for a younger age at maturity in their programs. Additionally, selection in fisheries for larger fish could also drive selection toward younger and smaller returning adult salmon. Ultimately selection in the direction of smaller and younger maturing salmon could result in the decline of recruitment for each generation, as fewer and fewer offspring are generated in each cohort.

Advancing our understanding of the genetic basis of age at maturity, would potentially allow hatchery programs to screen their brood stock for genetic markers that would favor later age at maturity. Current high-throughput sequencing and genetic marker discovery technologies in addition to advances in statistical approaches make it possible to identify regions of the genome associated with complex life history traits. Using these approaches we outline a project to sequence more than 400 Chinook salmon returns from Johnson Creek, a tributary of the Salmon River in Central Idaho, for more than 5,000 SNP markers distributed throughout the genome. The 400 target samples are subdivided into classes representing their origin (hatchery or wild), their sex (male or female) and their age at return (3, 4, or 5 years of age) in addition to their year of birth (Table 1). By identifying these categories, we can test different hypotheses about the effects and interactions of sex, origin, and cohort environment on the age at maturity, all factors suspected of contributing to the overall trait variation. The primary aim of this study, however, is to identify genetic markers closely linked to variation in age at maturity. Achieving this goal, we would contribute substantially to the early phases of dissecting the genetic basis of age at maturity in Chinook salmon, and potentially other salmonid fishes.

Table 1. A summary of the 408 Chinook salmon samples that will be used for SNP genotyping and genetic association mapping from Johnson Creek, ID. Samples are categorized based on their origin, either hatchery derived or wild (no hatchery influence), sex, age at maturity (3, 4, or 5 years of age), and Brood Year (BY, representing the year they were born).

| | Hatchery | | | | | | Wild | | | | | | |
| | Male | | | Female | | | Male | | | Female | | | |
BY/Age	3	4	5	3	4	5	3	4	5	3	4	5	Total
1998	15	8	31		9	10	10	9	18	1	7	10	128
2003	17	8	6	1	9	10	3	8	18		8	17	105
2004	17	8	1	1	9	6	22	11	4		10	10	99
2005	17	8	1		9	1	14	9	2		9	6	76
Total	66	32	39	2	36	27	49	37	42	1	34	43	408

Quantitative genetic analyses of both hatchery origin and wild origin pedigrees in Johnson Creek Chinook salmon indicate a moderate to high heritability in both age at sexual maturation (H^2=0.51-0.59) and in male "Jacking" (H^2=0.45-0.69). These results indicate that a large proportion of the variation in age at sexual maturation are controlled by a genetic component and that this life history can respond rapidly to natural and human mediated selection. Building upon this knowledge we designed a genome wide association study (GWAS) to identify regions of the genome and the distribution of genetic markers that are statistically associated with variation in age at maturity and male "Jacking". We genotyped 12,882 SNP markers from a panel of Johnson Creek samples described in Table 1. Among those 12,882 markers we identified 369 unique loci statistically associated with age at maturity or male "Jacking" in at least one of our

several models and a total of 547 significant association in all models. Preliminary alignment of our markers to a genetic linkage map indicates a genomewide distribution of associated loci, suggesting that several genes and genetic regions are associated with variation in age at maturity and male "Jacking".   Current efforts are aimed at identifying genes in the proximity of the significantly associated markers. Once candidate genes can be identified we can better understand the biological mechanisms of early male maturation ("Jacking") and shed light on how hatchery rearing might affect the expression of this life history trait.

Run-timing in summer/winter-run steelhead and spring/summer/fall-run Chinook salmon

We have initiated a study to investigate the genetic basis of run-timing of summer and winter run steelhead in the Klickitat River. Previous studies show that these winter and summer run fish interbreed extensively and are not genetically distinct when examined with neutral markers. Current results with ~15,000 quality filtered SNPs confirm previous signals of interbreeding between summer/winter runs, but preliminary Fst outlier analyses identified at least three SNPs show strong association with the run-time life-history. The pattern is extremely convincing and should serve for interesting content in a manuscript. The next steps will include additional RAD genotyping to fill in holes in the run distribution and more complete statistical analyses.

There is also an ongoing study to evaluate the genetic basis of run-timing in Chinook salmon from the Klickitat River. Approximately 15,000 SNPs have been genotyped for fish with run-timing phenotypes from adult fish returning to the trap at Lyle Falls. A subset of about 4,000 putatively neutral SNPs have been used to construct a phylogeny with Klickitat R. and Rangewide collections and the Klickitat R. population is intermediate but decidedly out on the Interior Stream-type branch as seen in previous studies. This signal of introgression between spring and fall lineages of Chinook salmon makes this an excellent system to investigate the genomic regions associated with run-timing. Next steps will involve using linkage maps to identify specific genomic regions that are associated with run-timing in this species.

Thermal adaptation in O. mykiss

For thermal adaptation, further testing is ongoing to more closely investigate genes that are associated with response to thermal stress including use of heart rate as a phenotype rather than survival. This study continues to use a combination of RAD-seq and RNA-seq to determine genetic associations with thermal adaptation in fish from desert and montane streams. Additionally, *O. mykiss* populations throughout the Columbia River Basin will be RAD sequenced to allow the opportunity to characterize genetic adaptation across a broader range of the species.