

University of Windsor

## Scholarship at UWindor

---

Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

---

2016

# IMPROVING SPECIES REINTRODUCTION THROUGH CONSERVATION GENOMICS

Xiaoping He  
*University of Windsor*

Follow this and additional works at: <https://scholar.uwindsor.ca/etd>

---

### Recommended Citation

He, Xiaoping, "IMPROVING SPECIES REINTRODUCTION THROUGH CONSERVATION GENOMICS" (2016).  
*Electronic Theses and Dissertations*. 5827.  
<https://scholar.uwindsor.ca/etd/5827>

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email ([scholarship@uwindsor.ca](mailto:scholarship@uwindsor.ca)) or by telephone at 519-253-3000ext. 3208.

# IMPROVING SPECIES REINTRODUCTION THROUGH CONSERVATION GENOMICS

By

Xiaoping He

A Dissertation

Submitted to the Faculty of Graduate Studies  
through the Great Lakes Institute for Environmental Research  
in Partial Fulfillment of the Requirements for  
the Degree of Doctor of Philosophy  
at the University of Windsor

Windsor, Ontario, Canada

2016

© 2016 Xiaoping He

# IMPROVING SPECIES REINTRODUCTION THROUGH CONSERVATION GENOMICS

by

Xiaoping He

APPROVED BY:

---

L. Bernatchez  
Laval University

---

T. Pitcher  
Department of Biological Sciences

---

K. Drouillard  
Great Lakes Institute for Environmental Research

---

G.D. Haffner  
Great Lakes Institute for Environmental Research

---

D. Heath, Advisor  
Great Lakes Institute for Environmental Research

September 6, 2016

## DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION

### I. Co-Authorship Declaration

I hereby declare that this thesis incorporates material that is result of joint research. All chapters were conducted under the supervision of Professor Daniel Heath. In Chapter 2, Mattias L. Johansson provided help with the manuscript revision. In Chapter 3, Chris C. Wilson was responsible for microsatellite genotyping; Kyle W. Wellband provided help with data analyses; Chris C. Wilson, Aimee Lee S. Houde and Bryan D. Neff provided facilities and took charge of fish rearing. In Chapter 4, Aimee Lee S. Houde provided help with data analyses; Trevor E. Pitcher was responsible for fish rearing. In Chapter 5, Aimee Lee S. Houde and Bryan D. Neff took charge of artificial stream tank setup and fish rearing. In Chapter 6, Subba Rao Chaganti provided advice for data analyses. In all cases, the key ideas, primary contributions, data analyses and interpretation, were performed by the author, Xiaoping He.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

### II. Declaration of Previous Publication

This thesis includes four original papers that have been previously published/submitted for publication in peer reviewed journals, as follows:

Thesis Chapter	Publication title /full citation	Publication status
Chapter 2	He X, Johansson ML, Heath DD (2016) Role of genomics and transcriptomics in selection of reintroduction source populations. <i>Conservation Biology</i> , <b>30</b> , 1010-1018.	Published
Chapter 3	He X, Wilson CC, Wellband KW, Houde ALS, Neff BD, Heath DD (2015) Transcriptional profiling of two Atlantic salmon strains: implications for reintroduction into Lake Ontario. <i>Conservation Genetics</i> , <b>16</b> , 277-287.	Published
Chapter 4	He X, Houde ALS, Pitcher TE, Heath DD. Genetic architecture of gene transcription in two Atlantic salmon ( <i>Salmo salar</i> ) populations.	Submitted
Chapter 6	He X, Chaganti SR, Heath DD. Population-specific responses to interspecific competition in the gut microbiota of two Atlantic salmon ( <i>Salmo salar</i> ) populations.	Submitted

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

## ABSTRACT

Reintroduction is the release of a species collected from captive or wild sources into its historical habitat where it has been locally extirpated with the aim to re-establish a self-sustaining population. Increasing pressures on global biodiversity caused by human activities has led to an upsurge in reintroductions in the last decades, but the reintroduction success rate is generally low. Populations can differ in reintroduction performance because of their genetic background which may limit their scope for adapting to novel environments as well as narrow their tolerance ranges for environmental stressors likely to be encountered in the initial acclimation phase of reintroduction. Thus, selecting an appropriate population is very important for conservation related applications including reintroduction. Atlantic salmon (*Salmo salar*) was extirpated in Lake Ontario by 1900s, and decades of reintroduction attempts have been largely unsuccessful. This dissertation focuses on two important reasons for the unsuccessful reintroduction of Atlantic salmon in Lake Ontario: inappropriate source population selection and stress caused by established non-native salmonids.

I explored population differences between two Atlantic salmon populations (LaHave and Sebago) and their responses to interspecific competition by characterization of gene expression and gut microbiota. The regulation of gene expression plays an important role in acclimation and adaptation. The gut microbial community mediates a variety of biological processes and can directly impact host fitness. In this dissertation, I addressed basic genetic (genetic components of gene expression variance), evolutionary (selection versus genetic drift on gene expression variance), and ecological (in response

to interspecific competition) theories of gene expression. I also addressed population differences in competitive ability and possible molecular mechanisms that mediated negative effects on Atlantic salmon caused by non-native competitors. I found that populations showed substantial differences in gene expression and genetic components of gene expression variance at rest state, and populations showed different response patterns to interspecific competition in gene expression and gut microbiota. The Sebago population is more suitable for reintroduction in Lake Ontario than the LaHave population. The results highlighted the fact that populations can possess different responses to biotic stressors despite not encountering the stressor during their past evolutionary history.



## ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor Dr. Daniel Heath for the incredible opportunity to come to Canada to pursue my PhD. I thank him for his brilliant guidance during the last five years. I would like to thank my committee members, Drs. Trevor Pitcher, Ken Drouillard and Doug Haffner for their advice and comments.

I would like to thank collaborators Drs. Chris Wilson and Bryan Neff for their support in this project. I would like to thank Bill Sloan, Scott Ferguson, Stephan Howailth, Andy Hunter, Andrew Smith and Craig Black for their invaluable assistance with fish tank construction, fish breeding, fish feeding and sample collection at the Ontario Ministry of Natural Resources and Forestry (OMNRF) Codrington Research Facility. I would like to thank Anne Kidd for microsatellite genotyping. Special thanks are owed to Dr. Aimee Lee Houde who played an indispensable role in coordinating with OMNRF staff and co-workers for this project.

I would like to thank all fellow labmates, friends, staff and faculty at GLIER for their assistance and support over the past five years, especially those who edited my writing and helped improve my English. Finally, I would like to thank my parents and my sister for their support throughout my life.

## TABLE OF CONTENTS

DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION .....	iii
ABSTRACT.....	vi
ACKNOWLEDGEMENTS.....	viii
LIST OF TABLES .....	xiii
LIST OF FIGURES .....	xv
CHAPTER 1 GENERAL INTRODUCTION .....	1
REINTRODUCTION .....	1
GENE EXPRESSION.....	3
ROLE OF GENE EXPRESSION IN ACCLIMATION AND ADAPTATION.....	4
METHODS TO QUANTIFY GENE EXPRESSION.....	5
ATLANTIC SALMON IN LAKE ONTARIO.....	6
THESIS OBJECTIVE AND STRUCTURE.....	8
REFERENCES .....	12
CHAPTER 2 ROLE OF GENOMICS AND TRANSCRIPTOMICS IN SELECTION OF REINTRODUCTION SOURCE POPULATIONS .....	17
INTRODUCTION .....	17
SOURCE POPULATION SELECTION.....	19
<i>Differences in response and tolerance of environmental stresses</i> .....	19
<i>Reintroduction performance variation</i> .....	20
FUNCTIONAL GENETIC VARIATION AND SINGLE NUCLEOTIDE POLYMORPHISMS.....	22
<i>Genetic similarity and genetic diversity</i> .....	22
<i>Single nucleotide polymorphism</i> .....	24
<i>Application of genome-wide SNPs to reintroduction</i> .....	25
GENE EXPRESSION AND APPLICATION OF TRANSCRIPTOMICS .....	26
<i>Gene expression and its importance</i> .....	26

<i>Population difference in stress response via gene transcription</i> .....	27
<i>Application of transcriptional profiling to reintroduction</i> .....	28
LIMITATIONS OF CONSERVATION GENOMICS AND TRANSCRIPTOMICS .....	29
CONCLUSIONS .....	31
REFERENCES .....	32
 CHAPTER 3 TRANSCRIPTIONAL PROFILING OF TWO ATLANTIC SALMON POPULATIONS: IMPLICATIONS FOR REINTRODUCTION INTO LAKE ONTARIO.....	39
INTRODUCTION .....	39
MATERIALS AND METHODS.....	43
<i>Atlantic salmon populations</i> .....	43
<i>Oligonucleotide microarray construction</i> .....	44
<i>RNA extraction, microarray hybridization and data preparation</i> .....	45
<i>Statistical analysis</i> .....	46
<i>Functional analysis</i> .....	48
<i>Microsatellite genotyping and <math>F_{ST}</math> estimation</i> .....	48
<i><math>P_{ST}</math> estimation</i> .....	49
RESULTS .....	50
<i>Tank effect on gene transcription</i> .....	50
<i>Differentially expressed genes between populations</i> .....	50
<i><math>F_{ST}</math> and <math>P_{ST}</math></i> .....	51
DISCUSSION .....	51
REFERENCES .....	57
 CHAPTER 4 GENETIC ARCHITECTURE OF GENE TRANSCRIPTION IN TWO ATLANTIC SALMON POPULATIONS .....	70
INTRODUCTION .....	70
MATERIALS AND METHODS.....	73
<i>Breeding design</i> .....	73
<i>RNA extraction and cDNA preparation</i> .....	74
<i>Gene selection</i> .....	74
<i>Quantitative real-time PCR</i> .....	75
<i>Statistical analysis</i> .....	77

RESULTS .....	78
<i>Population effects on gene transcription</i> .....	78
<i>Significant effects on gene transcription within each population</i> .....	79
<i>Additive genetic, non-additive genetic, and maternal effects</i> .....	79
DISCUSSION .....	81
CONCLUSIONS .....	85
REFERENCES .....	86
CHAPTER 5 TRANSCRIPTOME RESPONSE OF ATLANTIC SALMON TO COMPETITION WITH ECOLOGICALLY SIMILAR NON-NATIVE SPECIES .....	99
INTRODUCTION .....	99
MATERIALS AND METHODS .....	103
<i>Design and sampling</i> .....	103
<i>RNA isolation</i> .....	103
<i>Data analysis</i> .....	104
<i>Quantitative real-time PCR</i> .....	105
RESULTS .....	106
<i>Sequencing summary and reads mapping</i> .....	106
<i>Principle component analysis and distance heatmap</i> .....	107
<i>Gene expression differences</i> .....	108
<i>Comparison between gene expression level revealed by qRT-PCR and RNA-seq</i> .....	109
DISCUSSION .....	109
LIMITATIONS OF THIS STUDY .....	112
CONCLUSIONS .....	113
REFERENCES .....	114
CHAPTER 6 POPULATION-SPECIFIC RESPONSES TO INTERSPECIFIC COMPETITION IN THE GUT MICROBIOTA OF TWO ATLANTIC SALMON POPULATIONS .....	132
INTRODUCTION .....	132
MATERIALS AND METHODS .....	136
<i>Interspecific competition and sample collection</i> .....	136
<i>DNA extraction, PCR, and library preparation</i> .....	138
<i>Bioinformatic and statistical analysis</i> .....	139

RESULTS .....	141
<i>Summary of sequencing and core OTUs</i> .....	141
<i>Bacterial community composition</i> .....	142
<i>Effects of population and treatment on bacterial diversity</i> .....	143
<i>Population and treatment effects at the OTU level</i> .....	143
<i>Differences in beneficial bacteria and opportunistic pathogens</i> .....	144
DISCUSSION .....	145
CONCLUSIONS .....	150
REFERENCES .....	151
CHAPTER 7 GENERAL CONCLUSION .....	168
INTRODUCTION .....	168
CONTRIBUTIONS .....	170
<i>Population differences in gene expression</i> .....	170
<i>Population-specific responses to interspecific competition</i> .....	171
<i>Source population selection for reintroduction in Lake Ontario</i> .....	173
FUTURE DIRECTIONS .....	173
REFERENCES .....	176
APPENDIX A: REPRINT PERMISSIONS .....	179
VITA AUCTORIS .....	181

## LIST OF TABLES

### CHAPTER 3

**Table 3.1** List of genes showing significantly different transcription between the two replicate tanks for the Sebago population of Atlantic salmon (*Salmo salar*). ..... 62

**Table 3.2** List of genes showing significantly different transcription between the two replicate tanks for the LaHave population of Atlantic salmon (*Salmo salar*). ..... 63

**Table 3.3** List of genes showing significantly different transcription between the Sebago and LaHave populations of Atlantic salmon (*Salmo salar*). ..... 64

### CHAPTER 4

**Table 4.1** Quantitative real-time PCR primers and TaqMan probes for muscle function related genes (abbreviation in parentheses) used for Atlantic salmon (*Salmo salar*) genetic architecture analysis in two populations. .... 91

**Table 4.2** Summary of the genetic variance components of gene transcription for 22 genes associated with muscle function (arranged by putative function) in juvenile Atlantic salmon (*Salmo salar*) from two populations (LaHave and Sebago). ..... 93

**Table 4.3** Summary of genetic architecture for gene transcription of four functional categories across 22 muscle-function related genes for juvenile Atlantic salmon (*Salmo salar*) from two populations (LaHave and Sebago). ..... 95

**Supplementary Table S4.1** Model selection and population effects for the 22 analyzed genes in the two populations of Atlantic salmon (*Salmo salar*). ..... 98

### CHAPTER 5

**Supplementary Table S5.1** Primers used for qRT-PCR. .... 122

**Supplementary Table S5.2** Mapping summary of the 10 samples to Atlantic salmon (*Salmo salar*) draft genome. .... 123

**Supplementary Table S5.3** The expression of 10 genes which showed response to competition with all the three non-native salmonids for the LaHave Atlantic salmon (*Salmo salar*) population. .... 124

**Supplementary Table S5.4** The expression of nine genes which showed response to competition with all the three non-native salmonids for the Sebago Atlantic salmon (*Salmo salar*) population. .... 125

<b>Supplementary Table S5.5</b> The expression of 23 genes which showed response to competition with Chinook salmon ( <i>Oncorhynchus tshawytscha</i> ) for both Atlantic salmon ( <i>Salmo salar</i> ) populations. ....	126
---	-----

<b>Supplementary Table S5.6</b> The expression of 13 genes which showed response to competition with brown trout ( <i>Salmo trutta</i> ) for both Atlantic salmon ( <i>Salmo salar</i> ) populations. ....	127
--	-----

<b>Supplementary Table S5.7</b> The expression of 19 genes which showed response to competition with rainbow trout ( <i>Oncorhynchus mykiss</i> ) for both Atlantic salmon ( <i>Salmo salar</i> ) populations. ....	128
---	-----

## CHAPTER 6

<b>Table 6.1</b> Primer sequences used for Atlantic salmon ( <i>Salmo salar</i> ) gut microbiota characterization. ....	156
---	-----

## LIST OF FIGURES

### CHAPTER 2

**Figure 2.1** Application of conservation genomics and transcriptomics to source population selection for species reintroduction. Selected functional single nucleotide polymorphisms (SNPs) and genome-wide coding region SNPs (cSNP) can be used to measure functional genetic variation available for natural selection. Transcriptional profiling of candidate genes or whole transcriptome analysis can be used to quantify the population's scope for acclimation in response to environmental stress. .... 37

**Figure 2.2** Numbers of papers published that used DNA-based markers for conservation applications from 2005 to 2014 (all papers, all publications in which authors used microsatellite or SNP markers with a conservation application [{conservation} and {microsatellite or SNP}]; whole genome, only publications in which authors used a genome-wide marker approach [{conservation} and {whole genome}] followed by inspection for relevant studies). .... 38

### CHAPTER 3

**Figure 3.1** Gene transcription heatmap showing hierarchical clustering of the 21 differentially expressed genes between Sebago and LaHave Atlantic salmon (*Salmo salar*) populations. The rows represent different genes and the columns represent different tanks. The transcription level for each gene is the average  $\log_2$  transformed intensity value of fish from the same tank. .... 65

**Figure 3.2** Histograms of global  $P_{ST}$  for transcription of genes between two populations of juvenile Atlantic salmon (*Salmo salar*) arranged in increasing order. Panel a:  $P_{ST}$  values for all 271 analyzed genes. Panel b:  $P_{ST}$  values for genes which showed significantly different transcription levels between the two populations. The two horizontal dashed lines represent the upper and lower limits of the 99% confidence interval of  $F_{ST}$  based on microsatellite genotypes at eight loci. The solid line represents the mean  $F_{ST}$ . .... 66

**Supplementary Figure S3.1** Histograms of  $P$  values for the analysis of difference in gene transcription between two tanks of the Sebago Atlantic salmon (*Salmo salar*) population. Panel a:  $P$  value for each gene arranged in increasing order. Panel b: percentage of  $P$  value for each gene with the same order as (a) that is less than 0.05 during the 10 000 permutation. The solid line in (a) represents the statistical significance (0.05). .... 67

**Supplementary Figure S3.2** Histograms of  $P$  values for the analysis of difference in gene transcription between two tanks of the LaHave Atlantic salmon (*Salmo salar*) population. Panel a:  $P$  value for each gene arranged in increasing order. Panel b: percentage of  $P$  value for each gene with the same order as (a) that is less than 0.05



during the 10 000 permutation. The solid line in (a) represents the statistical significance (0.05). ..... 68

**Supplementary Figure S3.3** Histograms of  $P$  values for the analysis of difference in gene transcription between Sebago and LaHave Atlantic salmon (*Salmo salar*) populations. Panel a:  $P$  value for each gene arranged in increasing order. Panel b: percentage of  $P$  value for each gene with the same order as (a) that is less than 0.05 during the 10 000 permutation. The solid line in (a) represents the statistical significance (0.05). ..... 69

## CHAPTER 4

**Figure 4.1** Differences in gene transcription between two Atlantic salmon (*Salmo salar*) populations (LaHave and Sebago) measured using qRT-PCR. Gene names are described in Table 4.1. \*  $P < 0.05$ ; \*\* $P < 0.01$ . ..... 96

**Figure 4.2** Scatterplots showing the distribution of muscle function gene transcription variance components for two Atlantic salmon (*Salmo salar*) populations (LaHave and Sebago). Comparisons of variance components between populations: (a) additive genetic effects, (b) non-additive genetic effects, and (c) maternal effects. Each symbol represents one gene. Symbols are coded for the four different functional gene groups: circles (○) represent energy generation genes; squares (□) represent genes involved in energy regulation; triangles (△) represent growth genes; diamonds (◇) represent muscle contraction genes. The dashed line represents the expected 1:1 relationship (i.e.,  $y = x$ ) between the variance components of the two study populations. Points above the 1:1 line indicate higher values in the Sebago population and points below the dashed line indicate higher values in the LaHave population. .... 97

## CHAPTER 5

**Figure 5.1** Principal component analysis based on expression levels for 2 000 selected genes which exhibited the highest expression variation among samples for (a) all the 10 samples, (b) the LaHave population samples ( $n = 5$ ) and (c) the Sebago population samples ( $n = 5$ ). Treatment symbols: LaAS indicates LaHave Atlantic salmon reared alone; LaBT, LaCH, and LaRT indicate LaHave Atlantic salmon reared with one of the three species: brown trout, Chinook salmon, and rainbow trout, respectively; SeAS indicates Sebago Atlantic salmon reared alone; SeBT, SeCH, and SeRT indicate Sebago Atlantic salmon reared with one of the three species: brown trout, Chinook salmon, and rainbow trout, respectively. .... 119

**Figure 5.2** Heatmap of sample-to-sample distances based on expression levels of all the genes. Treatment symbols: AS indicates Atlantic salmon reared alone; BT, CH, and RT indicate the Atlantic salmon population reared with one of the three species: brown trout, Chinook salmon, and rainbow trout, respectively. .... 120

**Figure 5.3** Number of responding genes of Atlantic salmon (*Salmo salar*) to the presence of brown trout (BT), Chinook salmon (CH), and rainbow trout (RT). ..... 121

**Supplementary Figure S5.1** Venn diagram showing the overlap of differentially expressed genes in response to the presence of brown trout (BT), Chinook salmon (CH) and rainbow trout (RT) within the LaHave (a) and Sebago (b) Atlantic salmon populations. .... 129

**Supplementary Figure S5.2** Functional categorization of significantly different expressed genes in response to interspecific competition for the LaHave and Sebago Atlantic salmon (*Salmo Salar*) populations. For each population, responding genes were combined. GO terms containing less than five genes in both populations were not included. .... 130

**Supplementary Figure S5.3** Comparison of gene expression levels revealed by RNA-Seq and qRT-PCR for 14 genes. Expression of the 14 genes for RNA-Seq were log<sub>2</sub> transferred RPKM values from four samples: LaAS1 (LaHave Atlantic salmon reared alone sample 1), LaBT (LaHave Atlantic salmon reared with brown trout), SeAS1 (Sebago Atlantic salmon reared alone sample 1), and SeBT (Sebago Atlantic salmon reared with brown trout). qRT-PCR was used to quantify expression of the 14 genes for 12 individuals from the four treatments and then  $-\Delta C_T$  (calculated by  $\Delta C_{T\text{reference}} - \Delta C_{T\text{targeted}}$ ) values were averaged for each treatment. Spearman rank correlation analysis was conducted using the averaged  $-\Delta C_T$  for each treatment and log<sub>2</sub>RPKM. .... 131

## CHAPTER 6

**Figure 6.1** Relative abundance of bacterial classes for juvenile Atlantic salmon (*Salmo salar*) in response to interspecific competition. Displayed are classes with at least 0.1% relative abundance in one treatment. The “others” category includes unclassified sequences at the class level and the sum of all classes that occurred at less than 0.1% relative abundance. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. .... 157

**Figure 6.2** The 27 Operational Taxonomic Units (OTUs) that were significantly different in abundance between the two Atlantic salmon (*Salmo salar*) populations across all treatments. Taxonomic assignment beside each OTU identification number is the lowest taxonomic level obtained. .... 158

**Figure 6.3** The 13 OTUs showing difference among competition treatments in the Atlantic salmon (*Salmo salar*) LaHave population: OTU32, 116, 650 and 1271 (genus *Chryseobacterium*), OTU39 (genus *Flavobacterium*), OTU1517 (species *succinicans*), OTU1820, 2515 and 3189 (genus *Lactobacillus*), OTU2688 (order *Lactobacillales*), OTU3001 (family *Enterobacteriaceae*), OTU3181 (genus *Wautersiella*), OTU3900 (genus *Streptococcus*). Displayed are means  $\pm$  1SE for treatments. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native

salmonids. Different letters above the bars indicate significant differences assessed using Tukey's post hoc multiple comparisons ( $P < 0.05$ ). ..... 159

**Figure 6.4** The six lactic acid genera showing differences among treatments in the Atlantic salmon (*Salmo salar*) LaHave population. Displayed are means  $\pm$  1SE for treatments. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. Different letters above the bars indicate significant differences assessed using Tukey's post hoc multiple comparisons ( $P < 0.05$ ). ..... 160

**Supplementary Figure S6.1** Distribution of the number of high quality sequences generated by next generation sequencing of the amplified *16S rRNA* gene for gut microbiota characterization of Atlantic salmon (*Salmo salar*). The solid line represents the average number of sequence reads per sample. Each bar represents one sample..... 161

**Supplementary Figure S6.2** Distribution of the number of high quality sequences generated by next generation sequencing of the amplified *16S rRNA* gene. Displayed are means  $\pm$  1SD for treatments. The solid line represents the average number of sequence reads per sample. The dashed lines represent the average numbers of sequence reads in each of the two Atlantic salmon populations. Treatment symbols: AS indicates Atlantic salmon (*Salmo salar*) reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. .... 162

**Supplementary Figure S6.3** The proportion of 16S rRNA sequences accounted by the core OTUs shared by 70% of the gut microbiota samples in Atlantic salmon (*Salmo salar*) for each interspecific competition treatment. Displayed are means  $\pm$  1SD for treatments. The dashed lines represent the average proportion of reads for each of the two populations. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. .... 163

**Supplementary Figure S6.4** Relative abundance of bacterial phyla for the combined water samples (N=4) and the combined samples for all Atlantic salmon (*Salmo salar*) from the LaHave (N=82) and Sebago (N=96) populations. The bars show only phyla with at least 0.1% relative abundance in one of the two Atlantic salmon populations or combined water samples. The "others" category includes unclassified sequences at the phylum level and the sum of all phyla that occurred at less than 0.1% relative abundance. .... 164

**Supplementary Figure S6.5** Relative abundance of bacterial phyla for juvenile Atlantic salmon (*Salmo salar*) in response to interspecific competition. Displayed are phyla with at least 0.1% relative abundance in one treatment. The "others" category includes

unclassified sequences at the phylum level and the sum of all phyla that occurred at less than 0.1% relative abundance. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. .... 165

**Supplementary Figure S6.6** Relative abundance of bacterial genera for juvenile Atlantic salmon (*Salmo salar*) in response to interspecific competition. Displayed are genera with at least 0.1% relative abundance in one treatment. The “others” category includes unclassified sequences at the genus level and the sum of all genera that occurred at less than 0.1% relative abundance. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. .... 166

**Supplementary Figure S6.7** The relative abundance of potential pathogens in the gut microbiota of two Atlantic salmon (*Salmo salar*) populations. There were no significant differences among treatments within each population. The *Aeromonas* genus was undetectable in the data. Displayed are means  $\pm$  1SE for treatments. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. .... 167

## CHAPTER 1

### GENERAL INTRODUCTION

#### REINTRODUCTION

Reintroduction is the release of a species collected from captive or wild sources into its historical habitat where it has been locally extirpated with the aim to re-establish a self-sustaining population (IUCN/SSC 2013). Increasing pressures on global biodiversity caused by human activities has led to an upsurge in reintroductions in the last decades (Seddon *et al.* 2007; Butchart *et al.* 2010). For example, 218 animal species were reintroduced by 1998 and the number of reintroduced animal species increased to 489 by 2005 (Seddon *et al.* 2007). However, the success rates of reintroductions are generally low (Armstrong & Seddon 2008). Although Fischer & Lindenmayer (2000) observed that 26% of reintroductions of animal species are successful and Godefroid *et al.* (2011) found the reintroduction success rates for plant species were less than 20%, the actual number is probably much lower as successful reintroduction results are more likely to be published than failed reintroductions (Fischer & Lindenmayer 2000; Miller *et al.* 2014). For example, Miller *et al.* (2014) found the translocation success rate for herpetofauna in New Zealand in publications was 41.7%, while the actual success rate for all herpetofauna translocations in New Zealand was 8.1%.

Even after correcting the original causes of the extirpation, many factors can influence the outcome of a reintroduction. First, the quality of the receiving habitat, especially the abundance of predators, can directly affect the survival of released individuals (Wanless *et al.* 2002; Moorhouse *et al.* 2009). Second, the developmental

stage at release must be carefully selected because it not only affects the survival of reintroduced individuals, but can also affect the expense of the reintroduction project, as rearing to different developmental stages involves different labor and equipment costs for captive populations (Coghlan & Ringler 2004). Third, source population selection is a key factor to determine reintroduction success or failure (Schneider 2011; Forsman 2014). Population differences in stress response and tolerance of environmental fluctuations have been observed across species (Whitehead *et al.* 2010; Schoville *et al.* 2012; Wellband & Heath 2013), thus populations can exhibit variable survival and establishment performance after release. Populations also harbor different functional genetic diversity which is important for populations to evolve and persist in new environments (Lande & Shannon 1996; Montalvo *et al.* 1997). Finally, public attitudes and support contribute to the outcomes of reintroductions (Reading & Kellert 1993; Clark *et al.* 2002).

Successful reintroductions require both establishment and persistence of released individuals in the target habitat. *Establishment* refers to survival and reproduction of released individuals (Seddon *et al.* 2012), and *persistence* refers to the increase in numbers and density of reintroduced species in the recipient regions (Armstrong & Seddon 2008). The stages for successful reintroduction are similar to species invasion except that reintroduction requires that established population to persist in the target habitat while invasive species spread their living range (Armstrong & Seddon 2008; Blackburn *et al.* 2011). In both species reintroduction and invasion, establishment is a key stage in the success of the target species. Whether reintroduced and invasive individuals can survive in the receiving habitat depends on their acclimation capacity to stress imposed

by the local environment. In the long term, whether reintroduced species can persist or invasive species can spread may depend on their adaptive potential.

## GENE EXPRESSION

Gene transcription is the process whereby genetic information from a gene is used to synthesize RNA and is the first step in expressing functional products of protein-coding genes. Although gene expression includes gene transcription and translation, and regulatory processes post-transcription can affect expression level of proteins, quantification of gene transcription has been widely used to estimate expression levels of corresponding genes due to their theoretically and empirically high correlation. A variety of studies conducted in bacteria, yeast, and mammalian cells found that the correlation coefficients between mRNA abundance and protein levels ranged from 0.36 to 0.76 (Maier *et al.* 2009). Although the correlation between mRNA and protein levels varied from study to study, gene transcription is generally agreed to be a good proxy to estimate levels of gene expression (Li *et al.* 2014).

Gene expression plays an important role in phenotypic variation. It has been hypothesized that phenotypic differences among individuals, populations and even species are likely to be driven more by differential regulation of gene expression than changes in protein sequences (King & Wilson 1975). The expression of a gene can be measured using molecular biology techniques so that gene expression levels can be treated as an external phenotype and traditional quantitative genetic methods can be applied to map genomic regions that underlie expression levels of genes. Association and linkage studies have identified numerous expression quantitative trait loci (eQTL) in

model species. Based on the physical distance between eQTL and the expression levels of the genes the eQTL influence, eQTL can be divided into *cis*-acting eQTL and *trans*-acting eQTL. A *cis*-acting locus indicates that the DNA variation in the gene directly influences its transcription level, and a *trans*-acting locus is distant from the genes it influences, indicating that the genes in question are regulated by DNA variation at other genes or chromosome regions (Petretto *et al.* 2006). Studies also revealed that environmental factors and the interaction between genetics and environment contribute to gene expression variation as for other more traditional quantitative traits (Smith & Kruglyak 2008; López-Maury *et al.* 2008; Hodgins-Davis & Townsend 2009; Grishkevich & Yanai 2013).

#### ROLE OF GENE EXPRESSION IN ACCLIMATION AND ADAPTATION

As there is a close relationship between gene expression and fitness, and gene expression is determined by both genetic and environmental factors, the importance of gene expression in acclimation and adaptation is obvious. Gene expression responds to environmental changes and this response can be adaptive. For example, at the cellular level, the stabilization of hypoxia inducible factor 1 alpha subunit (HIF1 $\alpha$ ) and its binding to HIF1 $\beta$  in response to hypoxic conditions can regulate the expression of a variety of genes to maintain oxygen homeostasis (Wenger 2002). The production of different phenotypes by a single genotype in response to different environments (phenotypic plasticity) is thought to be processed by regulation of gene expression which can be achieved by up-/down-regulation of expression levels and/or selectively expressing alternatively spliced mRNAs (Schlichting & Smith 2002; Schulte 2004). Gene



expression variation is heritable and is closely related to phenotype; therefore, selection can act at the gene expression level. Evolution of phenotypes among populations and species likely depends more on variation in the regulation of gene expression than changes in protein sequences (King & Wilson 1975; Fraser 2013). The application of transcriptomic tools in various taxa further demonstrate that population differences in gene expression reflect different adaptations to temperature (Garvin *et al.* 2015), habitat (Huang *et al.* 2016), toxin tolerance (Whitehead *et al.* 2010), among others.

## METHODS TO QUANTIFY GENE EXPRESSION

There are a variety of methods used to quantify gene expression. Generally, mRNA is first isolated from targeted tissues or cells and then is reverse-transcribed to complementary DNA (cDNA), primarily because cDNA is much more stable for downstream analyses than mRNA. The abundance of cDNA for different genes is quantified as it reflects the transcription of corresponding genes in the samples from which mRNA is isolated. Quantitative real-time PCR (qRT-PCR) is the gold standard method to measure gene expression. qRT-PCR measures gene expression by monitoring the increase of amplified DNA during PCR as a reflection of the initial abundance of cDNA of the targeted genes. The monitoring of amplified DNA is assisted by fluorescent dye which is quantified at every PCR cycle. This method is limited in the number of gene expression assays possible simultaneously compared to other techniques. DNA microarray technology is a hybridization-based method whereby quantification of gene expression is also assisted by fluorescent dyes. A DNA microarray is a microscope glass slide on which gene-specific probes have been immobilized (Kammenga *et al.* 2007). To measure gene expression using a DNA microarray, mRNA is reverse-transcribed to

fluorescent-dye labeled cDNA which is then hybridized to the immobilized DNA probes on the slide. The signal of fluorescent dye for each hybridized probe is measured to estimate the quantity of mRNA (expression) of the corresponding genes (Duggan *et al.* 1999). Technical errors of DNA microarray analysis can be high due to issues caused by background noise and cross-hybridization (Zhang *et al.* 2005; Bengtsson & Bengtsson 2006). RNA-Seq is a sequencing-based method which applies high-throughput next generation sequencing technologies to sequence mRNA, and expression levels of genes are quantified based on sequence read counts via bioinformatic analysis. RNA-Seq can be used to profile transcriptomes for species without any prior gene sequence characterization. However, RNA-Seq is comparatively expensive, meaning that for most labs it is not affordable to quantify hundreds of individuals using this technique.

## ATLANTIC SALMON IN LAKE ONTARIO

Atlantic salmon (*Salmo salar*) was a native species abundant in most tributaries of Lake Ontario (Dunfield 1985). Declines of Atlantic salmon in Lake Ontario were observed in the 1840s (Parsons 1973), and this species was extirpated in the lake by the end of nineteenth century (Crawford 2001). The extirpation of Atlantic salmon in Lake Ontario was mainly caused by the construction of mill dams, which not only blocked Atlantic salmon migration to good-quality habitat for spawning, but also made them more easily caught by fishermen (Wright 1892; Parsons 1973). Other human-mediated activities, including deforestation, overharvesting and pollution, also contributed to the extirpation of Atlantic salmon in Lake Ontario (Wright 1892; Parsons 1973).

Attempts to reverse the decline of Atlantic salmon in Lake Ontario through intentional release date back to 1867 (Kerr 2006). Recent intensive reintroduction

attempts of Atlantic salmon into Lake Ontario by the governments of the Province of Ontario and New York State commenced in the 1980s (Crawford 2001), but those reintroductions have been unsuccessful (Stewart & Schaner 2002; COSEWIC 2006). While the environment in Lake Ontario has been greatly improved, the conditions in Lake Ontario may have changed substantially compared to historic conditions, including the establishment of exotic species, pollution by chlorinated organic compounds, and temperature increases (Beeton 2002). However, conditions in Lake Ontario tributaries currently are thought to be suitable for juvenile Atlantic salmon (Stanfield & Jones 2003).

There are a few possible reasons for the unsuccessful reintroduction of Atlantic salmon into Lake Ontario. First, the repeated unsuccessful reintroduction attempts using the LaHave source population implies that this population may not be a suitable source for Lake Ontario (Van Zwol *et al.* 2012). The LaHave population, originating from the LaHave River in Nova Scotia, is an anadromous population, while most historic Atlantic salmon in Lake Ontario are thought to be landlocked (Parsons 1973). Second, the establishment of non-native salmonids is thought to be a significant barrier to successful reintroduction of Atlantic salmon because intense competition can occur among the species due to niche overlap (Scott *et al.* 2003, 2005; Van Zwol *et al.* 2012; Houde *et al.* 2015a; b, 2016). Juvenile rainbow trout, brown trout and Atlantic salmon prefer riffle microhabitats (Hartman 1965; Morantz *et al.* 1987), and rainbow trout and brown trout are more aggressive than Atlantic salmon (Van Zwol *et al.* 2012). The presence of rainbow trout and brown trout caused detrimental effects on fitness-related traits of Atlantic salmon in both artificial stream tanks and natural streams (Houde *et al.* 2015a; b, 2016). Adult Chinook salmon have been reported to affect nest establishment and

decrease survival of mature Atlantic salmon during spawning in natural streams (Scott *et al.* 2003). Third, the establishment of introduced salmonid prey species (e.g. alewife and rainbow smelt) whose bodies contain high amounts of thiaminase, is another possible factor contributing to the failed reintroduction of Atlantic salmon (Ketola *et al.* 2000; Dimond & Smitka 2005). Consumption of these prey species can lead to thiamine deficiency in adult fish and thus cause high mortality (known as Early Mortality Syndrome) in their offspring by generation-transmitted thiamine deficiency (Fisher *et al.* 1996; Ketola *et al.* 2000; Coghlan & Ringler 2004).

## THESIS OBJECTIVE AND STRUCTURE

My thesis focuses on the establishment stage of reintroduction because population establishment is the prerequisite for population persistence and it is logically more important than persistence for species with high early mortality, such as fish. This thesis uses molecular biology techniques to explore Atlantic salmon population differences in gene expression and gut microbiota and their response to ecological challenges. The primary goal is to discern the implications for source population selection for Atlantic salmon reintroduction in Lake Ontario and other conservation and commercial applications. Pre-adaptation and adaptive responses of gene expression play an important role in an organisms' survival and establishment in new environments after release (Schlichting & Smith 2002; Schulte 2004). The microbial community in the gut involves in a variety biological processes of the host and the host physiology can in turn affect composition and diversity of gut microbiota (Sommer & Bäckhed 2013). Changes in gut microbiota in response to ecological challenges is an important factor to examine.

Population differences in gene expression and gut microbiota and their response to ecological challenges can provide information for source population selection for reintroduction and other conservation and commercial applications. Here, I investigated those areas in the chapters described below.

In Chapter 2, I reviewed population differences in responses to and tolerances of environmental stresses and changes due to their different evolutionary processes and genetic backgrounds. Population differences in adaptive potential and stress response is the fundamental basis for population differences in reintroduction performance difference. In this chapter, I propose the application of genome-wide functional genetic variation analyses to estimate adaptive potential and apply gene expression analyses to estimate acclimation and tolerance of stress for source population selection. Appropriate application of genomic and transcriptomic tools would promote more effective source population selection to increase reintroduction success.

In Chapter 3, I determined whether the difference in gene expression between the LaHave and Sebago Atlantic salmon populations is due to selection or genetic drift. To compare population differences in gene expression, I developed a custom oligo DNA microarray consisting of probes for 375 targeted genes which were selected either because of functional importance or responses to environmental stress and changes in previous studies. To determine evolutionary forces on gene expression variation, I calculated  $F_{ST}$  (a measurement of population divergence) between the two populations based on genotypes of neutral genetic markers and  $P_{ST}$  (a measurement of phenotypic differentiation) for each of the differentially expressed genes between the two

populations. I found 21 genes were differentially expressed between the two populations and the differences were likely driven by selection.

In Chapter 4, I conducted quantitative genetic analyses on genetic components of gene expression variance in two Atlantic salmon populations. I used a high-throughput qRT-PCR system to quantify expression of 22 genes in 426 Atlantic salmon from two populations (LaHave and Sebago) produced using 5×5 full factorial breeding designs. I analyzed population differences in gene expression and partitioned additive, non-additive and maternal effects of gene expression variance in these populations. I found that dams contributed more to gene expression variance than sires, but maternal effects were generally low. The average additive genetic effect of gene expression was smaller than previously reported for fitness-related traits in salmonids (Carlson & Seamons 2008). The results supported previous findings that gene expression is determined by genetic and environmental factors (Buckland 2004; Petretto *et al.* 2006). The results also indicated that gene expression evolves more slowly than fitness-related traits due to their small additive genetic effects.

In Chapter 5, I explored population differences in response to interspecific competition at gene expression level. I collected the spleens of Atlantic salmon from interspecific competition experiments between two Atlantic salmon populations (LaHave and Sebago) and three ecologically similar salmonids (Chinook salmon, rainbow trout and brown trout). I applied RNA-Seq to compare population differences in transcriptomic response to interspecific competition. I found population effects on gene expression were higher than interspecific competition effects. I also found population-specific responses to the same competitors. The results implied that rainbow trout may be the most stressful

competing species for the LaHave population and brown trout may be the most stressful competing species for the Sebago population, highlighting that transcriptomic tools can provide more detailed information than fitness-related traits in estimating stress response. The results indicated that RNA-Seq is a very effective tool to evaluate population differences in response to stress.

In Chapter 6, I measured population differences in response to interspecific competition in gut microbiota. I collected intestinal contents of Atlantic salmon from interspecific competition experiments between two Atlantic salmon populations (LaHave and Sebago) and four ecologically similar salmonids (Chinook salmon, coho salmon, rainbow trout and brown trout). I applied next generation sequencing to characterize gut microbiota for 178 Atlantic salmon and analyzed the effects of population and treatment on gut microbiota. I found that there were significant differences in the bacterial diversity and relative abundance of OTUs in the gut microbiota of the two populations. I found that the Sebago population had advantageous gut microbiota over the LaHave population and that treatment (competition) effects on gut bacteria were significant in the LaHave population but not in the Sebago population. The results also demonstrated that gut microbiota variation has the potential to be a good biomarker in selecting source populations for reintroduction and other conservation and commercial purposes.

## REFERENCES

- Armstrong DP, Seddon PJ (2008) Directions in reintroduction biology. *Trends in Ecology & Evolution*, **23**, 20–25.
- Beeton AM (2002) Large freshwater lakes: present state, trends, and future. *Environmental Conservation*, **29**, 21–38.
- Bengtsson A, Bengtsson H (2006) Microarray image analysis: background estimation using quantile and morphological filters. *BMC Bioinformatics*, **7**, 96.
- Blackburn TM, Pyšek P, Bacher S et al. (2011) A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, **26**, 333–339.
- Buckland PR (2004) Allele-specific gene expression differences in humans. *Human Molecular Genetics*, **13**, R255–260.
- Butchart SHM, Walpole M, Collen B et al. (2010) Global biodiversity: indicators of recent declines. *Science*, **328**, 1164–1168.
- Carlson SM, Seamons TR (2008) A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. *Evolutionary Applications*, **1**, 222–238.
- Clark JD, Huber D, Servheen C (2002) Bear reintroduction: lessons and challenges. *Ursus*, **13**, 335–345.
- Coghlan SM, Ringler NH (2004) A comparison of Atlantic salmon embryo and fry stocking in the Salmon River, New York. *North American Journal of Fisheries Management*, **24**, 1385–1397.
- COSEWIC (2006) Committee on the Status of Endangered Wildlife in Canada. *COSEWIC assessment and status report on the Atlantic salmon *Salmo salar* (Lake Ontario population) in Canada*. Committee on the Status of Endangered Wildlife in Canada, Ottawa, vii + 26 pp.
- Crawford SS (2001) Salmonine introductions to the Laurentian Great Lakes: an historical review and evaluation of ecological effects. *Canadian Special Publication of Fisheries and Aquatic Sciences* 132, 205 pp.
- Dimond P, Smitka J (2005) Evaluation of selected strains of Atlantic salmon as potential candidates for the restoration of Lake Ontario. *Trout Unlimited Canada Technical Report ON-012*, 41 pp.



- Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM (1999) Expression profiling using cDNA microarrays. *Nature Genetics*, **21**, 10–14.
- Dunfield RW (1985) The Atlantic salmon in the history of North America. *Canadian Special Publication of Fisheries and Aquatic Sciences* 80, 181 pp.
- Fischer J, Lindenmayer DB (2000) An assessment of the published results of animal relocations. *Biological Conservation*, **96**, 1–11.
- Fisher JP, Fitzsimons JD, Combs GF, Spitsbergen JM (1996) Naturally occurring thiamine deficiency causing reproductive failure in Finger Lakes Atlantic salmon and Great Lakes lake trout. *Transactions of the American Fisheries Society*, **125**, 167–178.
- Forsman A (2014) Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 302–307.
- Fraser HB (2013) Gene expression drives local adaptation in humans. *Genome Research*, **23**, 1089–1096.
- Garvin MR, Thorgaard GH, Narum SR (2015) Differential expression of genes that control respiration contribute to thermal adaptation in redband trout (*Oncorhynchus mykiss gairdneri*). *Genome Biology and Evolution*, **7**, 1404–1414.
- Godefroid S, Piazza C, Rossi G *et al.* (2011) How successful are plant species reintroductions? *Biological Conservation*, **144**, 672–682.
- Grishkevich V, Yanai I (2013) The genomic determinants of genotype  $\times$  environment interactions in gene expression. *Trends in Genetics*, **29**, 479–487.
- Hartman GF (1965) The role of behavior in the ecology and interaction of underyearling coho salmon (*Oncorhynchus kisutch*) and steelhead trout (*Salmo gairdneri*). *Journal of the Fisheries Research Board of Canada*, **22**, 1035–1081.
- Hodgins-Davis A, Townsend JP (2009) Evolving gene expression: from G to E to GxE. *Trends in Ecology & Evolution*, **24**, 649–658.
- Houde ALS, Smith AD, Wilson CC, Peres-Neto PR, Neff BD (2016) Competitive effects between rainbow trout and Atlantic salmon in natural and artificial streams. *Ecology of Freshwater Fish*, **25**, 248–260.
- Houde ALS, Wilson CC, Neff BD (2015a) Effects of competition with four nonnative salmonid species on Atlantic salmon from three populations. *Transactions of the American Fisheries Society*, **144**, 1081–1090.

- Houde ALS, Wilson CC, Neff BD (2015b) Competitive interactions among multiple non-native salmonids and two populations of Atlantic salmon. *Ecology of Freshwater Fish*, **24**, 44–55.
- Huang Y, Chain FJJ, Panchal M *et al.* (2016) Transcriptome profiling of immune tissues reveals habitat-specific gene expression between lake and river sticklebacks. *Molecular Ecology*, **25**, 943–958.
- IUCN/SSC (2013) *Guidelines for reintroductions and other conservation translocations. Version 1.0*. IUCN Species Survival Commission, Gland, Switzerland, viiii + 57 pp.
- Kammenga JE, Herman MA, Ouborg NJ, Johnson L, Breitling R (2007) Microarray challenges in ecology. *Trends in Ecology & Evolution*, **22**, 273–279.
- Kerr SJ (2006) *An historical review of fish culture, stocking and fish transfers in Ontario, 1865-2004*. Fish and Wildlife Branch, Ontario Ministry of Natural Resources, Peterborough, 222 pp.
- Ketola HG, Bowser PR, Wooster GA, Wedge LR, Hurst SS (2000) Effects of thiamine on reproduction of Atlantic salmon and a new hypothesis for their extirpation in Lake Ontario. *Transactions of the American Fisheries Society*, **129**, 607–612.
- King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science*, **188**, 107–116.
- Lande R, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, **50**, 434–437.
- Li JJ, Bickel PJ, Biggin MD (2014) System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ*, **2**, e270.
- López-Maury L, Marguerat S, Bähler J (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, **9**, 583–593.
- Maier T, Güell M, Serrano L (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Letters*, **583**, 3966–3973.
- Miller KA, Bell TP, Germano JM (2014) Understanding publication bias in reintroduction biology by assessing translocations of New Zealand’s herpetofauna. *Conservation Biology*, **28**, 1045–1056.
- Montalvo AM, Williams SL, Rice KJ *et al.* (1997) Restoration biology: a population biology perspective. *Restoration Ecology*, **5**, 277–290.

- Moorhouse TP, Gelling M, Macdonald DW (2009) Effects of habitat quality upon reintroduction success in water voles: evidence from a replicated experiment. *Biological Conservation*, **142**, 53–60.
- Morantz DL, Sweeney RK, Shirvell CS, Longard DA (1987) Selection of microhabitat in summer by juvenile Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **44**, 120–129.
- Parsons JW (1973) History of salmon in the Great Lakes, 1850-1970. *Technical Papers of the Bureau of Sport Fisheries and Wildlife*, Washington D.C., 80 pp.
- Petretto E, Mangion J, Dickens NJ *et al.* (2006) Heritability and tissue specificity of expression quantitative trait loci. *PLoS Genetics*, **2**, e172.
- Reading RP, Kellert SR (1993) Attitudes toward a proposed reintroduction of black-footed ferrets (*Mustela nigripes*). *Conservation Biology*, **7**, 569–580.
- Schlichting CD, Smith H (2002) Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, **16**, 189–211.
- Schneider J (2011) Review of reintroduction of Atlantic salmon (*Salmo salar*) in tributaries of the Rhine River in the German Federal States of Rhineland-Palatinate and Hesse. *Journal of Applied Ichthyology*, **27**, 24–32.
- Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS (2012) Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. *BMC Evolutionary Biology*, **12**, 170.
- Schulte PM (2004) Changes in gene expression as biochemical adaptations to environmental change: a tribute to Peter Hochachka. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, **139**, 519–529.
- Scott RJ, Noakes DLG, Beamish FWH, Carl LM (2003) Chinook salmon impede Atlantic salmon conservation in Lake Ontario. *Ecology of Freshwater Fish*, **12**, 66–73.
- Scott RJ, Poos MS, Noakes DLG, Beamish FWH (2005) Effects of exotic salmonids on juvenile Atlantic salmon behaviour. *Ecology of Freshwater Fish*, **14**, 283–288.
- Seddon PJ, Armstrong DP, Maloney RF (2007) Developing the science of reintroduction biology. *Conservation Biology*, **21**, 303–312.
- Seddon PJ, Strauss WM, Innes J (2012) Animal translocations: What are they and why do we do them? In: *Reintroduction Biology* (eds Ewen JG, Armstrong DP, Parker KA, Seddon PJ), pp. 1–32. John Wiley & Sons, Ltd.

- Smith EN, Kruglyak L (2008) Gene–environment interaction in yeast gene expression. *PLoS Biology*, **6**, e83.
- Sommer F, Bäckhed F (2013) The gut microbiota--masters of host development and physiology. *Nature Reviews. Microbiology*, **11**, 227–238.
- Stanfield L, Jones ML (2003) Factors influencing rearing success of Atlantic salmon stocked as fry and parr in Lake Ontario tributaries. *North American Journal of Fisheries Management*, **23**, 1175–1183.
- Stewart TJ, Schaner T (2002) Lake Ontario salmonid introductions 1970 to 1999: stocking, fishery and fish community influences. In: *Lake Ontario Fish Communities and Fisheries: 2001 Annual Report of the Lake Ontario Management Unit*. Queen's Printer for Ontario, Picton, pp. 12.1-12.10.
- Van Zwol JA, Neff BD, Wilson CC (2012) The effect of competition among three salmonids on dominance and growth during the juvenile life stage. *Ecology of Freshwater Fish*, **21**, 533–540.
- Wanless RM, Cunningham J, Hockey PAR *et al.* (2002) The success of a soft-release reintroduction of the flightless Aldabra rail (*Dryolimnas [cuvieri] aldabranus*) on Aldabra Atoll, Seychelles. *Biological Conservation*, **107**, 203–210.
- Wellband KW, Heath DD (2013) Environmental associations with gene transcription in Babine Lake rainbow trout: evidence for local adaptation. *Ecology and Evolution*, **3**, 1194–1208.
- Wenger RH (2002) Cellular adaptation to hypoxia: O<sub>2</sub>-sensing protein hydroxylases, hypoxia-inducible transcription factors, and O<sub>2</sub>-regulated gene expression. *The FASEB Journal*, **16**, 1151–1162.
- Whitehead A, Triant DA, Champlin D, Nacci D (2010) Comparative transcriptomics implicates mechanisms of evolved pollution tolerance in a killifish population. *Molecular Ecology*, **19**, 5186–5203.
- Wright RR (1892) Preliminary report on the fish and fisheries of Ontario. *Ontario Game and Fish Commission* 26148: 419-476.
- Zhang J, Finney RP, Clifford RJ, Derr LK, Buetow KH (2005) Detecting false expression signals in high-density oligonucleotide arrays by an in silico approach. *Genomics*, **85**, 297–308.

## CHAPTER 2

### ROLE OF GENOMICS AND TRANSCRIPTOMICS IN SELECTION OF REINTRODUCTION SOURCE POPULATIONS<sup>1</sup>

#### INTRODUCTION

Reintroduction is the intentional release of a species into its historical range where it has become extirpated and is distinct from reinforcement, where organisms are translocated to existing populations of the same species (IUCN/SSC 2013). Because of sharp declines in global biodiversity caused by human activities (Butchart *et al.* 2010), reintroduction has become an important conservation tool and is likely to increase in its application as extirpation rates increase (IUCN/SSC 2013). By 1998, 218 animal species had been reintroduced, and that number doubled by 2005 (Fischer & Lindenmayer 2000; Seddon *et al.* 2007). Although the number of reintroduction projects is increasing worldwide, project success rates are generally low (Armstrong & Seddon 2008). Forty-nine percent of animal reintroductions with known outcomes can be considered successful (Fischer & Lindenmayer 2000), and the success rate for freshwater fish is as much as 58% (Cochran-Biederman *et al.* 2015). The real success rate may be much lower due to publication biases that drive higher publication rates for successful reintroduction reports relative to failed reintroductions and the high percentage of reintroductions with uncertain outcomes (Fischer & Lindenmayer 2000; Miller *et al.* 2014).

---

<sup>1</sup> He X, Johansson ML, Heath DD (2016) Role of genomics and transcriptomics in selection of reintroduction source populations. *Conservation Biology*, **30**, 1010-1018

Given good quality receiving habitat, the selection of an appropriate source population is an essential and critical factor affecting reintroduction success (Schneider 2011; IUCN/SSC 2013; Forsman 2014). Populations can differ in reintroduction potential due to their genetic architecture (genome content and epistatic interactions), which may limit their scope for adapting to novel environments and narrow their tolerance range for environmental stressors encountered in the initial acclimation phase of reintroduction. In general, genetic architecture is not only closely related to fitness in a locally adapted population but also a key factor in determining whether organisms can survive and thrive in novel or changing environments (Lande & Shannon 1996; Ouborg *et al.* 2010). The importance of the genetic architecture of source populations is reflected in the IUCN Guidelines for Reintroductions (IUCN/SSC 2013) which state that if a translocation consists of many individuals with high genetic diversity, then source genetic architecture may not limit reintroduction success.

Conservation genomics is a new field that applies genomic technologies to address conservation questions (Ouborg *et al.* 2010; Hoffmann *et al.* 2015). Some applications of conservation genomics simply increase the power of traditional population-genetics approaches by using more DNA markers to cover more of the genome (Kohn *et al.* 2006). However, conservation genomics and transcriptomics can also address more challenging long-standing issues in conservation, including quantifying fitness-related genetic variation, measuring how environmental stress affects gene activity, and determining the molecular mechanisms of tolerance to environmental fluctuation and pollutants (Ouborg *et al.* 2010; Whitehead *et al.* 2010; Harrison *et al.* 2014). First, I argue that the selection of source population is critical for successful

species reintroduction and must go beyond simple inbreeding avoidance (i.e., maximize neutral genetic diversity). I then explain how genomic tools can be used to facilitate selection of the optimal source populations based on two criteria: maximizing functional genetic variation to foster adaptive potential and maximizing potential adaptive plasticity to foster acclimation, or breadth of tolerance. I argue that the appropriate application of conservation genomics and transcriptomics has the potential to dramatically improve the success rate of reintroduction, a critical tool for maintaining biodiversity in the face of rapid environmental change.

## SOURCE POPULATION SELECTION

### *Differences in response and tolerance of environmental stresses*

Because of divergent evolutionary processes, populations of the same species that differ in morphology, behavior, life history, and physiology occur in almost all kingdoms. Population differences exist in static traits and occur in response to environmental change or stressors. Populations of *Populus davidiana* subjected to different levels of drought stress differ in their strategies to survive drought (Zhang *et al.* 2004). Côte *et al.* (2012) compared embryo incubation time, body size, and survival of four Atlantic salmon (*Salmo salar*) populations under normoxic and hypoxic conditions and observed significant population by environment interactions and parental effects for all traits. Southern populations of killifish (*Fundulus heteroclitus*) exhibit higher plasma cortisol response to acute and chronic handling stress than northern populations (DeKoning *et al.* 2004). These examples highlight the broad range of responses to environmental stresses among presumably locally adapted populations and, hence, differences in the adaptive potential of these populations to new environments after reintroduction.

Populations also differ in their tolerance of environmental stress and resistance to pathogens. Even after growing in a common garden for 30 years, populations of white ash (*Fraxinus americana*) differ in their cold tolerance, growth, and survival (Marchin *et al.* 2008). Fanguie *et al.* (2006) compared thermal tolerance in six populations of the common killifish (*Fundulus heteroclitus*) and found that the three southern populations have a higher critical thermal maximum than the three northern populations and that the latter had a lower critical thermal minimum. In Sweden, southern populations of the common lizard (*Lacerta vivipara*) are more resistant to viral eye disease (higher survival after infection) than northern populations (Uller *et al.* 2003). In general, differences in environmental stress tolerance and pathogen resistance are reported across taxa, indicating a high likelihood that potential source populations differ in response traits that can affect their expected survival and performance after reintroduction.

#### *Reintroduction performance variation*

Successful reintroductions require that the released organisms be able to establish and persist in the target habitat. *Establishment* refers to survival and successful reproduction (Seddon *et al.* 2012), and *persistence* refers to increased numbers and density of reintroduced species in the target habitat (Armstrong & Seddon 2008). When individuals are reintroduced, they face novel selective pressures imposed by the new environment. Thus, the survival of reintroduced individuals depends on how closely their phenotype matches locally favored phenotypes (Ghalambor *et al.* 2007) or, more generally, how well their phenotypes match the current conditions in the target habitat. Ideally, researchers should evaluate potential source populations for matching habitat



characteristics and the associated genetic architecture and phenotypic variation to choose the best possible source for reintroduction (Sork *et al.* 2013). However, such an exhaustive evaluation of candidate source populations is often not feasible.

Generally, there are two main mechanisms by which organisms can cope with environmental stress: phenotypic plasticity and genetic adaptation (Hansen *et al.* 2012; Harrisson *et al.* 2014). Phenotypic plasticity is the ability of individual genotypes to exhibit alternative phenotypes in response to different environments (Pigliucci *et al.* 2006). Genetic adaptation occurs when the genetic background of individuals within a population changes over time to maximize fitness in the new environment (Hendry *et al.* 2011; Hansen *et al.* 2012). In the short term, reintroduced individuals may survive through environmental acclimation via phenotypic plasticity (Schlichting & Smith 2002; Ghalambor *et al.* 2007). In the long term, genetic adaptation may be the key mechanism for introduced populations to survive and thrive (Hansen *et al.* 2012).

Because populations show substantial phenotypic differences at rest and in response to stress, it is not surprising that they may exhibit variation in reintroduction performance. White Storks (*Ciconia ciconia*) originating from their native northeastern European range have higher reproductive success and more offspring per pair than White Storks originating from North Africa (Olsson 2007). Schneider (2011) tested several Atlantic salmon populations for reintroduction into the Rhine River, and concluded that only the Swedish Ätran population was suitable for reintroduction because its spawning time matches the spawning time of the original population of the Rhine River. The variation in performance among potential source populations is likely explained by differences in their genetic architecture shaped by historic evolutionary pressures.

Whether reintroduced organisms can survive and thrive in new environments depends on their acclimation and adaptation potential, both determined by the evolutionary history of the source population. It is thus important to characterize potential source populations based on their current functional genetic variation (adaptive potential) and their scope for response to ecologically relevant stresses (acclimation). However, characterizing functional trait differences can be technically and logistically difficult, and neutral genetic variation, even with large numbers of loci, may not accurately reflect genome-wide functional genetic variation. I devised a new paradigm for reintroduction source population selection: conservation genomics and transcriptomics (Figure 2.1).

## FUNCTIONAL GENETIC VARIATION AND SINGLE NUCLEOTIDE POLYMORPHISMS

### *Genetic similarity and genetic diversity*

The genetic variation of source populations can significantly impact the outcome of reintroduction efforts. Two main aspects of genetic variation must be taken into account when selecting a source population: genetic similarity and genetic diversity. Ideally, a donor population that is genetically similar to the historical (extirpated) population in the targeted habitat should be selected for reintroduction. After environmental remediation or natural habitat recovery, the species could reestablish after reintroduction due to preexisting adaptations to the target habitat. For example, aurora trout (*Salvelinus fontinalis timagamiensis*) were extirpated from Whitepine and Whirligig Lakes in the 1960s due to acidification (Snucins *et al.* 1995). In 1990, when the lakes recovered, captive aurora trout originating from wild individuals collected from the same lakes before extirpation were reintroduced and natural reproduction was observed in

Whirligig Lake in 1992 (Snucins *et al.* 1995). However, in most cases individuals from the original population are not available, and there are usually no DNA samples from the original population that can be used to compare genetic similarity between the original and potential source populations (Schwartz 2005). Furthermore, although a historic habitat may appear to be restored, it is likely that current conditions are not the same; thus, even the original genetic stock may fail to reestablish. In most situations, one must select from extant populations of unknown genetic relatedness to the original population.

Using genetic diversity as a criterion for selecting source populations for reintroduction (Earnhardt 1999; IUCN/SSC 2013) ensures sufficient genetic variation for natural selection to act upon in the novel environment, maximizing adaptive potential (Lande & Shannon 1996). Avoiding low genetic diversity resulting from past genetic bottlenecks and inbreeding is also important for reintroduction success (Frankham 1995). This is the theoretical basis for using measures of genetic variation as surrogates for fitness in conservation (e.g., Reed & Frankham 2001). Although high genetic diversity is important for population fitness, not all genetic variation is related to fitness.

Applications of neutral genetic markers in conservation have increased dramatically (Ouborg *et al.* 2010), but the vast majority of those applications rely on small numbers of loci that may not reflect genome-wide diversity (Figure 2.2). Furthermore, although one may assume that neutral marker diversity is correlated with functional genetic variation, this may not be correct (Hedrick 2001; Reed & Frankham 2001). Ideally, genome-wide coverage based on functional marker loci should be used to achieve more complex conservation goals than inbreeding assessment and genetic isolation quantification.

### *Single nucleotide polymorphism*

Single nucleotide polymorphisms (SNPs) are DNA sequence differences at a single nucleotide where the less frequent allele of the polymorphism is 1% or higher (Vignal *et al.* 2002). Generally, SNPs are biallelic and distributed throughout the genome with high density (e.g., 1 SNP every 116 base pairs in the genome of channel catfish [*Ictalurus punctatus*] [Sun *et al.* 2014]). SNPs can be located in the coding region of genes, in introns, and between the genes (Jukema & Agema 2001). Coding region SNPs (cSNPs) can be further divided into synonymous and nonsynonymous SNPs.

Nonsynonymous SNPs are associated with changes in amino acid sequence and are thus most likely to represent functional genetic variation, although synonymous SNPs may be in linkage disequilibrium with unrecognized functional variation.

Variation in phenotype among and within populations is partially explained by variation in DNA sequence (Botstein & Risch 2003), and understanding of how variation at specific gene loci affects phenotype is growing rapidly. Thus, identifying genetic markers (e.g., SNPs) for variation in specific traits that are critical for reintroduction success will facilitate effective selection of source populations and individuals for reintroduction. For example, Johnston *et al.* (2014) used 4,353 SNPs to conduct a genome-wide association study between SNPs and the age at which Atlantic salmon return from the sea to spawn and identified 10 SNPs associated with maturation age. Such studies show the power of genomic approaches to identify functional DNA markers that can be used to evaluate source populations for reintroduction. However, despite the promise of conservation genomics for improving the selection of source populations for successful reintroductions and a rapidly decreasing cost to characterize genome-wide

SNPs, few such studies have been reported and no studies report reintroduction success. Thus, despite the rapid growth of conservation genetics and the broad acceptance of the concepts of conservation genomics, to date genomics and transcriptomics have rarely been used to assist in species reintroduction efforts (Figure 2.2).

#### *Application of genome-wide SNPs to reintroduction*

Morin *et al.* (2004) proposes SNPs replace microsatellite markers for applications in conservation genetics. Genotyping of genome-wide SNPs has been reported for many species, including farm animals (Muir *et al.* 2008; Petersen *et al.* 2013), fish (Willing *et al.* 2010; Jones *et al.* 2012), and plants (Grattapaglia *et al.* 2011; Plomion *et al.* 2014). Considering the importance of functional genetic variation and logistical convenience of genotyping genome-wide SNPs, I propose the use of SNP-based genome scans to estimate genetic diversity for selection of reintroduction source populations (Figure 2.1).

I further suggest that two types of SNP genome scans are useful for reintroduction efforts: nonsynonymous SNPs and SNPs associated with fitness-related traits. Nonsynonymous SNPs change protein sequences and thus may reflect variation in protein function. The SNPs already known to be associated with fitness-related traits may have either direct effects on phenotype or be in linkage disequilibrium with genetic variation underlying phenotypic variation. In either case, the estimation of genetic diversity at such SNP markers can effectively provide estimates of functional genetic variation (Figure 2.1). My proposed approach to known functional SNP scanning is appropriate for species with abundant genomic information and well-characterized gene function. However, many species of conservation concern have little genomic information and the genetic bases for fitness-related traits are barely studied. Therefore, I

propose genotyping anonymous genome-wide cSNPs to estimate genetic diversity for species with poor genomic characterization. Although not all cSNPs are functionally important, they are much more likely to be associated with functional genetic variation than known neutral DNA markers. Once appropriate functional SNPs are identified, candidate populations with the highest observed heterozygosity should be selected as source populations. This focus on maximizing heterozygosity in cSNPs will not only maximize functional genetic variation for functional traits and hence the likelihood of reintroduction success through adaptation but also provide useful information regardless of source population size because it provides objective functional criteria for choosing the source population with maximal adaptive potential for reintroduction.

## GENE EXPRESSION AND APPLICATION OF TRANSCRIPTOMICS

### *Gene expression and its importance*

Gene expression variation is the primary mechanism that leads to phenotypic variation within and among populations (Rifkin *et al.* 2003; Storey *et al.* 2007) because gene expression is influenced by both genetic and environmental factors (Petretto *et al.* 2006; López-Maury *et al.* 2008) specifically through regulation of expression, selective expression, or silencing of genes (Schlichting & Smith 2002). Therefore, variation in gene expression is a physiological process and an important adaptive mechanism allowing organisms to respond to novel habitats or environmental change or stress (Schulte 2004).

*Population difference in stress response via gene transcription*

With the rapid development of transcriptomic tools, it has become simpler to simultaneously quantify transcription at thousands of gene loci, even in species with few genomic resources available. Differentially expressed genes (at rest or in response to a challenge) represent potentially adaptive genetic variation among populations (Whitehead & Crawford 2006; Larsen *et al.* 2007; Giger *et al.* 2008; He *et al.* 2015). Thus, transcriptional profiling can lead to quantitative estimates of relative environmental stress response among populations. Transcriptional patterns in six wild rainbow trout populations differ in response to temperature and immune challenges, despite that the populations are separated by <250 km (Wellband & Heath 2013). Whitehead *et al.* (2010) compared transcriptome differences in response to polychlorinated biphenyl exposure between naturally tolerant and sensitive killifish populations and found that low expression of genes involved in the aryl hydrocarbon receptor signaling pathway may be a mechanism of pollution tolerance in killifish. Although examples such as these underscore a large body of literature that demonstrates that populations generally respond to environmental stressors differently at the gene transcription level, the application of that knowledge to the selection of source populations for reintroductions is practically nonexistent. The only published example of transcriptomics applied for reintroduction purposes focuses on the reintroduction of extirpated Atlantic salmon into Lake Ontario, Canada. Using a custom microarray, He *et al.* (2015) found significant gene transcription differences at 21 genes between two possible source populations, demonstrated that differences are likely the result of selection, and recommended one source population for reintroduction based on those differences. Differences in gene expression patterns

underlie the mechanisms of differential tolerance to environmental stress, and transcriptional profiling is thus an ideal, but underutilized, tool for selecting source populations for reintroduction.

#### *Application of transcriptional profiling to reintroduction*

Because variation in gene expression is directly linked to phenotypic variation, gene expression is potentially a powerful tool for the prediction of phenotypes (Oellrich *et al.* 2014). For example, Tung *et al.* (2012) compared gene expression among 10 rhesus macaque (*Macaca mulatta*) groups with different social status and found that the identified differentially expressed genes could be used to predict social rank class with 80% accuracy. Miller *et al.* (2011) applied a nonlethal biopsy method to collect tissues and used genome-wide gene transcription in wild migrating Sockeye salmon (*Oncorhynchus nerka*) to identify a genomic signature that could be used to predict migration and spawning success. Although this kind of application of transcriptomics is still in its infancy and few examples exist in the literature, those that do exist demonstrate the possibilities.

Because of the plastic nature of gene expression and its important role in the adaptive response to environmental stressors (acclimation), I propose the application of transcriptional profiling to quantify acclimation potential among potential source populations (Figure 2.1). Because many species of conservation concern occur only in small, highly fragmented populations, this focus on flexibility avoids the problem that small populations may be dominated by genetic drift and thus may no longer be locally adapted (Willi *et al.* 2006; Leimu & Fischer 2008). For species with well-characterized stress-response genes, transcriptional profiling could target a selected panel of candidate



genes associated with coping with ecologically relevant stressors and environmental fluctuation. For species whose stress-response genes are poorly characterized, whole transcriptome analysis should be used to quantify acclimation potential. In such cases, ecologically relevant and physical environmental challenges would have to be applied to individuals from the candidate populations and their gene transcription profile assessed before and after the challenge. Based on their transcriptional response, one can evaluate and predict population performance upon reintroduction in two ways. First, for cases where one knows a gene's function in mitigating environmental stress, one selects populations exhibiting adaptive responses. Second, for cases where one does not know whether upregulation or downregulation of a gene is beneficial, one selects populations with high plasticity.

## LIMITATIONS OF CONSERVATION GENOMICS AND TRANSCRIPTOMICS

Species reintroduction will become more common as habitats are altered and lost due to human activities and climate change. However, reintroduction efforts are costly, potentially environmentally risky, and subject to complex regulatory requirements (IUCN/SSC 2013). Thus, reintroductions are only feasible when reintroduction is central to species conservation and the species is of high priority. I contend that conservation genomics and transcriptomics are realistic possibilities for improving the likelihood of reintroduction success in key high-risk situations. Although the costs of genome-wide cSNP application and transcriptome profiling have come down substantially, it is still a major hurdle for often financially limited conservation efforts. I do not propose that genomics or transcriptomics should be the first line of response in a reintroduction effort; rather, they are promising tools for which the cost is dropping rapidly. More importantly,

the effectiveness of the conservation genomics and transcriptomics applications I propose have not yet been tested in any reintroduction that I am aware of; thus, no empirical data exist that shows it materially improves reintroduction success. Ideally, translocation experiments in controlled systems should be used to test whether functional genetic variation is a good predictor for long-term introduction success or whether transcriptional profiling can predict short-term acclimation and survival.

Because it is likely not feasible to apply conservation genomic and transcriptomic methods to reintroduction efforts for endangered species, I propose an empirical test of the application of genomic and transcriptomic techniques in an artificial reintroduction experiment. Using short-lived, genomically well-characterized species as models, groups could be introduced into controlled environments that range in environmental conditions such that they represent benign to potentially lethal levels of environmental stress. The putative source populations would be characterized as having either high or low levels of functional genetic variation and adaptive or nonadaptive transcriptional response to the environmental stressors in the artificial target habitats. Groups from the contrasting source populations would be introduced into the range of target habitats, and population performance would be monitored as reproduction in the short term and as population size and habitat-use expansion in the medium term. I predict that high functional genetic variation and high plasticity in transcriptional scope will drive increased short- and medium-term performance. Although the primary purpose of this essay is to make the argument that conservation genomics and transcriptomics has great promise and should be explored as a valuable tool in addressing the growing biodiversity conservation crisis, I cannot yet provide concrete evidence for its value in conservation efforts. However, the

growing understanding of how the genome and transcriptome is shaped by interactions with the environment provides compelling evidence for conservation genomics and transcriptomics as emerging and valuable tools for effectively managing the world's biodiversity.

## CONCLUSIONS

Functional SNP genotyping and transcriptional profiling are potentially powerful tools for reintroduction in particular and conservation in general. Conservation genomics and transcriptomics can not only answer long-standing questions in conservation biology but also provide important applications in reintroduction biology, specifically in selecting appropriate source populations. Characterizing genome-wide functional SNPs can provide quantitative estimates of fitness-related genetic variation and transcriptional profiling can provide data on how individuals respond to environmental stresses. Such data would have immediate practical applications in reintroduction biology as metrics for source population selection. Because successful reintroduction requires both short-term acclimation and long-term adaptation to the targeted habitat, I strongly urge conservation professionals to consider using functional SNP scans to measure genetic diversity and transcriptional profiling to measure the response of candidate and anonymous genes as part of the selection process for source populations for reintroduction. When based on genomic and transcriptomic measurements of adaptation and acclimation, the selection of source population will be more effective and will increase the success rate of reintroductions globally.

## REFERENCES

- Armstrong DP, Seddon PJ (2008) Directions in reintroduction biology. *Trends in Ecology & Evolution*, **23**, 20–25.
- Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics*, **33 Suppl**, 228–237.
- Butchart SHM, Walpole M, Collen B *et al.* (2010) Global biodiversity: indicators of recent declines. *Science*, **328**, 1164–1168.
- Cochran-Biederman JL, Wyman KE, French WE, Loppnow GL (2015) Identifying correlates of success and failure of native freshwater fish reintroductions. *Conservation Biology*, **29**, 175–186.
- Côte J, Roussel JM, Le Cam S, Bal G, Evanno G (2012) Population differences in response to hypoxic stress in Atlantic salmon. *Journal of Evolutionary Biology*, **25**, 2596–2606.
- DeKoning ABL, Picard DJ, Bond SR, Schulte PM (2004) Stress and interpopulation variation in glycolytic enzyme activity and expression in a teleost fish *Fundulus heteroclitus*. *Physiological and Biochemical Zoology*, **77**, 18–26.
- Earnhardt JM (1999) Reintroduction programmes: genetic trade-offs for populations. *Animal Conservation*, **2**, 279–286.
- Fangue NA, Hofmeister M, Schulte PM (2006) Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *The Journal of Experimental Biology*, **209**, 2859–2872.
- Fischer J, Lindenmayer DB (2000) An assessment of the published results of animal relocations. *Biological Conservation*, **96**, 1–11.
- Forsman A (2014) Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 302–307.
- Frankham R (1995) Conservation genetics. *Annual Review of Genetics*, **29**, 305–327.
- Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, **21**, 394–407.

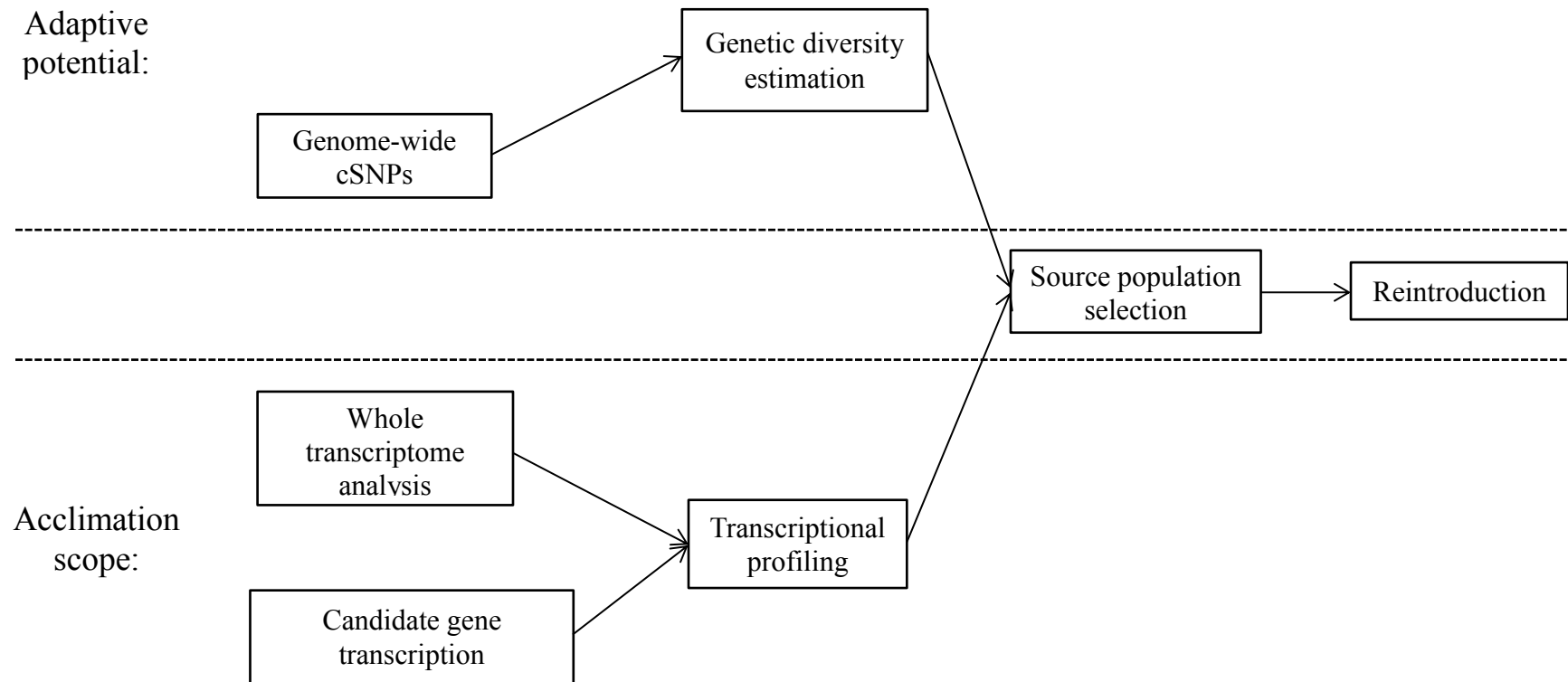
- Giger T, Excoffier L, Amstutz U *et al.* (2008) Population transcriptomics of life-history variation in the genus *Salmo*. *Molecular Ecology*, **17**, 3095–3108.
- Grattapaglia D, Silva-Junior OB, Kirst M *et al.* (2011) High-throughput SNP genotyping in the highly heterozygous genome of *Eucalyptus*: assay success, polymorphism and transferability across species. *BMC Plant Biology*, **11**, 65.
- Hansen MM, Olivieri I, Waller DM, Nielsen EE, GeM Working Group (2012) Monitoring adaptive genetic responses to environmental change. *Molecular Ecology*, **21**, 1311–1329.
- Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P (2014) Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, **7**, 1008–1025.
- He X, Wilson CC, Wellband KW *et al.* (2015) Transcriptional profiling of two Atlantic salmon strains: implications for reintroduction into Lake Ontario. *Conservation Genetics*, **16**, 277–287.
- Hedrick PW (2001) Conservation genetics: Where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.
- Hendry AP, Kinnison MT, Heino M *et al.* (2011) Evolutionary principles and their practical application. *Evolutionary Applications*, **4**, 159–183.
- Hoffmann A, Griffin P, Dillon S *et al.* (2015) A framework for incorporating evolutionary genomics into biodiversity conservation and management. *Climate Change Responses*, **2**, 1–24.
- IUCN/SSC (2013) Guidelines for reintroductions and other conservation translocations. Version 1.0. IUCN Species Survival Commission, Gland, Switzerland.
- Johnston SE, Orell P, Pritchard VL *et al.* (2014) Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (*Salmo salar*). *Molecular Ecology*, **23**, 3452–3468.
- Jones FC, Chan YF, Schmutz J *et al.* (2012) A genome-wide SNP genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current Biology*, **22**, 83–90.
- Jukema JW, Agema WRP (2001) The pharmacogenetics of atherosclerosis. In: *Cardiovascular Genetics for Clinicians*. (eds Doevendans PA, Wilde AAM), pp. 89–100. Springer Netherlands.

- Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. *Trends in Ecology & Evolution*, **21**, 629–637.
- Lande R, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, **50**, 434–437.
- Larsen PF, Nielsen EE, Williams TD *et al.* (2007) Adaptive differences in gene expression in European flounder (*Platichthys flesus*). *Molecular Ecology*, **16**, 4674–4683.
- Leimu R, Fischer M (2008) A meta-analysis of local adaptation in plants. *PloS One*, **3**, e4010.
- López-Maury L, Marguerat S, Bähler J (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, **9**, 583–593.
- Marchin RM, Sage EL, Ward JK (2008) Population-level variation of *Fraxinus americana* (white ash) is influenced by precipitation differences across the native range. *Tree Physiology*, **28**, 151–159.
- Miller KA, Bell TP, Germano JM (2014) Understanding publication bias in reintroduction biology by assessing translocations of New Zealand’s herpetofauna. *Conservation Biology*, **28**, 1045–1056.
- Miller KM, Li S, Kaukinen KH *et al.* (2011) Genomic signatures predict migration and spawning failure in wild Canadian salmon. *Science*, **331**, 214–217.
- Morin PA, Luikart G, Wayne RK, the SNP workshop group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216.
- Muir WM, Wong GK-S, Zhang Y *et al.* (2008) Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17312–17317.
- Oellrich A, Sanger Mouse Genetics Project, Smedley D (2014) Linking tissues to phenotypes using gene expression profiles. *Database*, bau017.
- Olsson O (2007) Genetic origin and success of reintroduced white storks. *Conservation Biology*, **21**, 1196–1206.
- Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma RK, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. *Trends in Genetics*, **26**, 177–187.

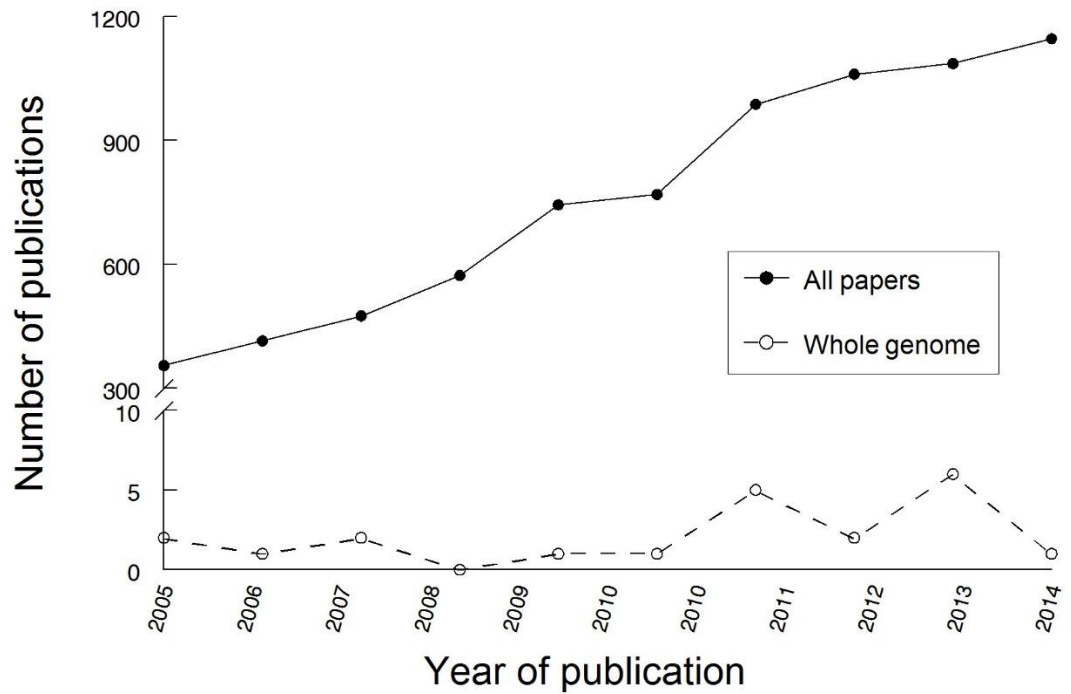
- Petersen JL, Mickelson JR, Cothran EG *et al.* (2013) Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PloS One*, **8**, e54997.
- Petretto E, Mangion J, Dickens NJ *et al.* (2006) Heritability and tissue specificity of expression quantitative trait loci. *PLoS Genetics*, **2**, e172.
- Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. *The Journal of Experimental Biology*, **209**, 2362–2367.
- Plomion C, Chancerel E, Endelman J *et al.* (2014) Genome-wide distribution of genetic diversity and linkage disequilibrium in a mass-selected population of maritime pine. *BMC Genomics*, **15**, 171.
- Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution*, **55**, 1095–1103.
- Rifkin SA, Kim J, White KP (2003) Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nature Genetics*, **33**, 138–144.
- Schlichting CD, Smith H (2002) Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, **16**, 189–211.
- Schneider J (2011) Review of reintroduction of Atlantic salmon (*Salmo salar*) in tributaries of the Rhine River in the German Federal States of Rhineland-Palatinate and Hesse. *Journal of Applied Ichthyology*, **27**, 24–32.
- Schulte PM (2004) Changes in gene expression as biochemical adaptations to environmental change: a tribute to Peter Hochachka. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, **139**, 519–529.
- Schwartz MK; (2005) Guidelines on the use of molecular genetics in reintroduction programs. In: *The EU LIFE Nature Projects to Guidelines for the Reintroduction of Threatened Species*, pp. 51–58. Caramanico Terme, Italy. Available from [http://www.fs.fed.us/rm/pubs\\_other/rmrs\\_2005\\_schwartz\\_m001.pdf](http://www.fs.fed.us/rm/pubs_other/rmrs_2005_schwartz_m001.pdf) (accessed July 2014).
- Seddon PJ, Armstrong DP, Maloney RF (2007) Developing the science of reintroduction biology. *Conservation Biology*, **21**, 303–312.
- Seddon PJ, Strauss WM, Innes J (2012) Animal translocations: What are they and why do we do them? In: *Reintroduction Biology* (eds Ewen JG, Armstrong DP, Parker KA, Seddon PJ), pp. 1–32. John Wiley & Sons, Ltd.
- Snucins EJ, Gunn JM, Keller W (1995) Restoration of the Aurora trout to its acid-damaged native habitat. *Conservation Biology*, **9**, 1307–1311.

- Sork VL, Aitken SN, Dyer RJ *et al.* (2013) Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, **9**, 901–911.
- Storey JD, Madeoy J, Strout JL *et al.* (2007) Gene-expression variation within and among human populations. *American Journal of Human Genetics*, **80**, 502–509.
- Sun L, Liu S, Wang R *et al.* (2014) Identification and analysis of genome-wide SNPs provide insight into signatures of selection and domestication in channel catfish (*Ictalurus punctatus*). *PloS One*, **9**, e109666.
- Tung J, Barreiro LB, Johnson ZP *et al.* (2012) Social environment is associated with gene regulatory variation in the rhesus macaque immune system. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 6490–6495.
- Uller T, Olsson M, Madsen T (2003) Family and population effects on disease resistance in a reptile. *Heredity*, **91**, 112–116.
- Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics, Selection, Evolution*, **34**, 275–305.
- Wellband KW, Heath DD (2013) Environmental associations with gene transcription in Babine Lake rainbow trout: evidence for local adaptation. *Ecology and Evolution*, **3**, 1194–1208.
- Whitehead A, Crawford DL (2006) Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 5425–5430.
- Whitehead A, Triant DA, Champlin D, Nacci D (2010) Comparative transcriptomics implicates mechanisms of evolved pollution tolerance in a killifish population. *Molecular Ecology*, **19**, 5186–5203.
- Willi Y, Buskirk JV, Hoffmann AA (2006) Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 433–458.
- Willing E-M, Bentzen P, van Oosterhout C *et al.* (2010) Genome-wide single nucleotide polymorphisms reveal population history and adaptive divergence in wild guppies. *Molecular Ecology*, **19**, 968–984.
- Zhang X, Zang R, Li C (2004) Population differences in physiological and morphological adaptations of *Populus davidiana* seedlings in response to progressive drought stress. *Plant Science*, **166**, 791–797.





**Figure 2.1** Application of conservation genomics and transcriptomics to source population selection for species reintroduction. Selected functional single nucleotide polymorphisms (SNPs) and genome-wide coding region SNPs (cSNP) can be used to measure functional genetic variation available for natural selection. Transcriptional profiling of candidate genes or whole transcriptome analysis can be used to quantify the population's scope for acclimation in response to environmental stress.



**Figure 2.2** Numbers of papers published that used DNA-based markers for conservation applications from 2005 to 2014 (all papers, all publications in which authors used microsatellite or SNP markers with a conservation application [{conservation} and {microsatellite or SNP}]; whole genome, only publications in which authors used a genome-wide marker approach [{conservation} and {whole genome}] followed by inspection for relevant studies).

## CHAPTER 3

# TRANSCRIPTIONAL PROFILING OF TWO ATLANTIC SALMON POPULATIONS: IMPLICATIONS FOR REINTRODUCTION INTO LAKE ONTARIO<sup>2</sup>

## INTRODUCTION

The conservation genetics paradigm is that small and isolated populations are subject to loss of genetic diversity and increased levels of homozygosity that in turn lead to increased likelihood of extirpation (Frankham *et al.* 2002; Ouborg *et al.* 2010). Loss of genetic diversity is thought to reduce individual fitness and affect the ability of a population to adaptively respond to a changing environment (Frankham 2003; Spielman *et al.* 2004). Therefore, conserving genetic diversity is often an important component of conservation plans and efforts. However, it is not clear if these efforts actually conserve functional genetic variation.

There are three forms of genetic variation in populations: neutral, deleterious, and adaptive (Hedrick 2001). Adaptive genetic variation is variation in coding or regulatory genes that have the potential to increase fitness (Hedrick 2001; Garcia de Leaniz *et al.* 2007). Thus, using functional genetic variation to address issues in conservation should be advantageous in comparison with neutral genetic variation. One form of functional genetic variation is gene expression variation, that is, the process whereby genomic variation is converted into phenotypic variation. One mechanism by which phenotypic

---

<sup>2</sup> He X, Wilson CC, Wellband KW, Houde ALS, Neff BD, Heath DD (2015) Transcriptional profiling of two Atlantic salmon strains: implications for reintroduction into Lake Ontario. *Conservation Genetics*, **16**, 277-287.

variation can arise from a single genotype is regulation in gene expression: this can be either sensitive up- or down-regulation or, more simply, on-off control (Schlichting & Smith 2002). Many studies have suggested that variation in gene expression among populations can be adaptive (Oleksiak *et al.* 2002; Whitehead & Crawford 2006; Larsen *et al.* 2007; Luca *et al.* 2009; Wellband & Heath 2013) and thus could be used to address conservation issues. Compared to neutral genetic variation, measures of variation in gene expression could provide more relevant information as it reflects the activity of functional genes. For example, Giger *et al.* (2006) profiled gene transcription using DNA microarrays and genotyped microsatellite loci in juveniles from six brown trout (*Salmo trutta*) populations and found that gene expression variation among populations was more affected by population life history (migratory or residential) than by their genetic distance based on neutral DNA markers.

Population differences in gene expression have been demonstrated in several species over the past decade. Some studies suggest that among-population gene expression variation is much higher than within-population variation (Townsend *et al.* 2003; Hutter *et al.* 2008), whereas others have found the opposite pattern (Oleksiak *et al.* 2002; Storey *et al.* 2007). However, those studies all indicated that variation in gene expression is an important source of variance for adaptation, and thus ultimately, evolution.

More recently, gene expression comparisons among populations have been applied to conservation. For example, gene transcription comparisons have revealed that introgression can result in changes in gene transcription profiles in both Atlantic salmon (*S. salar*) and brook charr (*Salvelinus fontinalis*), which may result in loss of local

adaptation (Roberge *et al.* 2008; Lamaze *et al.* 2013). Also, analyses of population differences in gene expression versus neutral DNA in 12 Atlantic salmon populations showed that gene transcription can be used to identify conservation units and has many advantages over the more traditional, neutral markers (Hansen 2010; Vandersteen Tymchuk *et al.* 2010). Pedersen *et al.* (2005) compared the expression of *heat shock protein 70 (Hsp70)* between inbred and outbred lines of *Drosophila melanogaster* to illustrate mechanisms of inbreeding depression and found that there was a significant negative correlation between transcription level of *Hsp70* and resistance to heat stress. Miller *et al.* (2011) collected gill tissue using nonlethal biopsy method from wild-caught Sockeye salmon (*Oncorhynchus nerka*) and identified a set of genes whose transcription can be used to predict migration and spawning success both in fresh water and in the ocean.

Atlantic salmon are broadly distributed in North America and Europe but they have declined or been extirpated in many rivers over the last 200 years (Parrish *et al.* 1998). Atlantic salmon was once an abundant fish in Lake Ontario (Ontario, Canada), but had disappeared by 1900, mainly because of habitat degradation (Crawford 2001). Because of its economic, ecological, and cultural value, there have been increasing efforts to reintroduce Atlantic salmon into Lake Ontario over the past three decades (Dimond & Smitka 2005); however, those reintroduction attempts have been unsuccessful. Potential explanations for the reintroduction failure of Atlantic salmon into Lake Ontario include environmental changes of Lake Ontario during the past years, such as establishment of non-native prey fish and non-native competitors (Coghlan & Ringler 2004; Scott *et al.* 2005; Houde *et al.* 2015), and perhaps inappropriate Atlantic salmon

population used for the reintroduction (Van Zwol *et al.* 2012). Although the Lake Ontario habitat has improved, selecting appropriate source populations is a crucial step for Atlantic salmon reintroduction, as populations can differ in their adaptive potential and environmental tolerances. To address possible limitations in candidate populations for reintroduction into Lake Ontario, I explored the genetic background (neutral and transcriptional) of two Atlantic salmon populations: LaHave and Sebago. LaHave is an anadromous population which originates from the LaHave River, Nova Scotia. The LaHave population has been used for reintroduction into Lake Ontario for many years and it was successfully reintroduced into Trout Lake, Ontario (Dimond & Smitka 2005). Sebago is a landlocked population from Sebago Lake, Maine. This population has a relatively large body size compared to other populations and it performed well in the Lake Champlain reintroduction where salmonid competitors (rainbow trout and brown trout) existed (Dimond & Smitka 2005; Van Zwol *et al.* 2012). In this study, I constructed a custom oligonucleotide microarray to compare gene transcription at selected known-function genes in gill tissue between the two source populations. I then calculated  $F_{ST}$  based on microsatellite genotypes and  $P_{ST}$  based on gene transcription levels, and used the  $F_{ST}$ - $P_{ST}$  comparison to identify selection versus genetic drift effects on the genes differentially expressed between the two populations. The results demonstrate how populations differ in gene expression and the evolutionary forces underlying those differences.

## MATERIALS AND METHODS

### *Atlantic salmon populations*

Two Atlantic salmon populations were provided by the Ontario Ministry of Natural Resources (OMNR): LaHave and Sebago. The LaHave population was from broodstock that has been in captivity for three generations. The Sebago population was derived from hatchery-bred fish that were released and recaptured as returning mature fish in Sebago Lake. Eggs and milt were collected from the recaptured adults and brought to Ontario in 2006 and the Sebago population was reared in captivity to be used as broodstock. Eggs from both populations were fertilized on November 4, 2010 at OMNR Harwood Fish Culture Station, Harwood, Ontario, and then reared at the OMNR Codrington Research Facility, Codrington, Ontario. Detailed information about the families and rearing environment is provided in Houde *et al.* (2013). Briefly, for each population, a full factorial cross was conducted using five males and five females to yield 25 full-sib families. Fertilized eggs were incubated in vertical stack incubators followed by rearing in tanks. The fish were transferred to artificial stream tanks in September 2011 as juvenile fry. Each artificial stream tank consisted of a riffle and a pool. More details about the juvenile salmon and artificial stream tank construction are described in Houde *et al.* (2015). In each tank, there were a total of 32 Atlantic salmon from eight families of one population with equal numbers (four fish) per family. Each stream tank was replicated once such that fish from each population were reared in two stream tanks (for a total of four stream tanks). In July 2012, after 10 months in the artificial stream tanks, eight fish from each tank were euthanized by overdose of tricaine methanesulfonate solution and gill tissue was collected and preserved in RNAlater. I chose gill tissue

because of its vital function in respiration, osmoregulation, nitrogen balance and disease resistance, and its fast response to environmental stressors (e.g. toxins and pathogens) relative to other organs (Campos-Perez *et al.* 2000; Evans *et al.* 2005). At the time of sampling, body mass ranged from 12.76-45.71 g with an average body mass ( $\pm$ SE) of  $24.24 \pm 1.68$  g. There were no significant differences in mean body mass between populations or among tanks.

#### *Oligonucleotide microarray construction*

To compare transcriptional differences between the two populations, I developed a custom oligonucleotide microarray. Custom microarrays have a few advantages compared to commercial microarrays: relatively low price, higher replication and more focused set of genes. My custom microarray consisted of probes for 380 different genes: 375 genes from Atlantic salmon and five control genes from *Arabidopsis thaliana*. Of the 375 genes, 277 genes were selected because of their functional importance and their mRNA sequences were obtained from the consortium for Genomics Research on All Salmon Project website (<http://web.uvic.ca/grasp/microarray>). The sequences of the other 98 genes were downloaded from Nucleotide database of NCBI (<http://www.ncbi.nlm.nih.gov>) and most of those genes had been reported to show transcriptional response to environmental factors. The five plant genes (*isoflavonoid reductase*, *psbP*, *psbW*, *salt-stress induced tonoplast intrinsic protein*, *ribulose 1.5-biphoshate carboxylase small subunit*) were used as negative control and their sequences were downloaded from NCBI. The probes were designed by using OligoArray 2.0 (Rouillard *et al.* 2003). The length of probes ranged from 45 to 55 nucleotides and  $T_m$  ranged from 82 to 88°C. A list and detailed information for the selected genes and probe



sequences are presented in Supplementary Table S3.1 ([https://static-content.springer.com/esm/art%3A10.1007%2Fs10592-014-0657-1/MediaObjects/10592\\_2014\\_657\\_MOESM2\\_ESM.xlsx](https://static-content.springer.com/esm/art%3A10.1007%2Fs10592-014-0657-1/MediaObjects/10592_2014_657_MOESM2_ESM.xlsx)). The oligonucleotide probes were printed on poly-L-Lysine coated slides (Thermo Scientific, USA) using a SpotArray 24 Microarray Printing System (PerkinElmer, Canada). On each slide, the probes were printed in three blocks (top, middle, and bottom) and each probe was printed three times adjacently within each block. Thus each probe was printed nine times on every slide. After printing, the probes were cross-linked to the slides by ultraviolet irradiation. The microarrays used in this experiment were printed in two batches and the potential batch effect caused by different printing was taken into account in the data analysis.

#### *RNA extraction, microarray hybridization and data preparation*

Gill tissue was placed in 2 mL tubes containing 1 mL TRIzol (Invitrogen, USA) and approximately 400  $\mu$ L of 1.0 mm diameter glass beads (BioSpec Products, USA). The tissue samples were homogenized at speed 6 for 40 seconds in a Thermo Savant FastPrep homogenizer (Lab Recyclers Inc., USA). Total RNA isolation followed the manufacturer's instructions ([http://tools.lifetechnologies.com/content/sfs/manuals/trizol\\_reagent.pdf](http://tools.lifetechnologies.com/content/sfs/manuals/trizol_reagent.pdf)). The concentration and purity of RNA was measured by spectrophotometry on a NanoVue spectrophotometer (GE Healthcare Bio-Science Corp, USA), and the quality of RNA was assessed by running 1  $\mu$ g of total RNA on a 1% agarose gel. Single colour microarray measurement was performed for this experiment using Array 50™ Cy3 Kit (Genisphere Inc., Hatfield, USA). Detailed protocols for reverse transcription, cDNA concentration, hybridization and washing are given in the Array 50™ Cy3 Kit's instruction

([http://genisphere.com/sites/default/files/pdf/Array50\\_Jan2011.pdf](http://genisphere.com/sites/default/files/pdf/Array50_Jan2011.pdf)). Briefly, 15-20 µg total RNA was reverse transcribed using SuperScript® II Reverse Transcriptase (Invitrogen) and RT primer (5' - TTCTCGTGTTCCGTTTGTACTCTAAGGTGGA–T(17)- 3'). The cDNA was concentrated and hybridized to microarrays for 12 h at 43°C. The slides were subsequently washed using 2X SSC with 0.2% SDS, 2X SSC and 0.2X SSC, separately. The slides were centrifuged immediately for two minutes at 1000 RPM to dry. The slides were then hybridized with Cy3-labeled fluorescent DNA dendrimer for 2.5 h at 43°C. The slides were washed and dried again as described above, then the slides were immediately scanned using a ScanArray Express microarray scanner (PerkinElmer, Canada) with the laser at 90% power and photo-multiplier tube (PMT) gain at 75%.

The scanned images were analyzed using ScanArray Express Microarray Analysis System software version 4.0 (PerkinElmer, Canada). Each spot was quantified using the adaptive circle method and the three blocks on each slide were quantified separately. After quantification, the data were background corrected and normalized using the *limma* package of R (Smyth 2005). First, the spots which failed to meet the quality criteria were filtered out. Then, “normexp” algorithm with an offset of 50 was used for background correction. After that, “quantile” normalization method was used to conduct between-array normalization. Finally, genes that had expression data in less than 70% of the spots across all samples were removed. The intensity of fluorescence for the remaining genes was log<sub>2</sub> transformed for statistical analyses.

### *Statistical analysis*

To detect artificial stream tank effects on gene expression, I analyzed the data population by population as tanks were nested in each population in this experiment. The

analysis was conducted using the *lme4* package of R (Bates & Maechler 2009) with the following model:

$$Y_{ijklm} = \mu + T_i + Ba_j + I_k + Bl_{l(k)} + e_{ijklm} \quad (1)$$

where  $Y_{ijklm}$  is the  $\log_2$  transformed normalized intensity value for each spot;  $\mu$  is the average value;  $T_i$  is the  $i^{\text{th}}$  effect of tank;  $Ba_j$  is the  $j^{\text{th}}$  effect of printing batch;  $Bl_{l(k)}$  is the  $l^{\text{th}}$  block effect (position on the array) which is nested within the  $k^{\text{th}}$  individual (fish) and  $e_{ijklm}$  is the random residual. Significance of expression differences between replicate tanks (within population) was determined using a likelihood ratio test between two models: one with and one without the tank effect included. Due to the complicated nature of the mixed-effects model and the dependency structure of the genes, resampling based False Discovery Rate (FDR) corrections are not supported. In lieu of these, I calculated the probability of detecting a  $P$ -value as extreme as the one I observed by randomly permuting the data 10 000 times and refitting the model for each gene to determine its significance under a completely null hypothesis. I report the probability of detecting the gene as significant as the number of times the permuted  $P$ -values were more extreme (less) than the observed  $P$ -value for that gene divided by the total number of permutations.

To test for gene transcription differences between the two Atlantic salmon populations (Sebago and LaHave), I used the following model:

$$Y_{ijklmn} = \mu + P_i + T_j + Ba_k + I_l + Bl_{m(l)} + e_{ijklmn} \quad (2)$$

where  $Y_{ijklmn}$  is the  $\log_2$  transformed normalized intensity value for each spot;  $\mu$  is the average value;  $P_i$  is the  $i^{\text{th}}$  effect of population;  $T_j$  is the  $j^{\text{th}}$  effect of tank;  $Ba_k$  is the  $k^{\text{th}}$  effect of array printing batch;  $Bl_{m(l)}$  is the  $m^{\text{th}}$  effect of block (position on the array) which

is nested within  $l^{\text{th}}$  individual (fish) and  $e_{ijklmn}$  is the random residual. The significance of the population effect was determined using a likelihood ratio test between two models: one with and one without the population effect included. I followed the same methodology detailed above to calculate the probability of false discovery for the population effect analyses.

### *Functional analysis*

The functions of those differentially expressed genes were analyzed in NCBI (<http://www.ncbi.nlm.nih.gov>) and the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Huang *et al.* 2009). The transcriptional level for the differentially expressed genes was averaged within each tank and then the data was used to construct a heat map using TM4 software (Saeed *et al.* 2003).

### *Microsatellite genotyping and $F_{ST}$ estimation*

To calculate neutral  $F_{ST}$  between the two stains, I used microsatellite genotype data for a total of 520 fish, of which 219 were collected from the Sebago population in Sebago Lake, and 301 were collected from LaHave population in Harwood Hatchery. Each fish was genotyped at eight microsatellite loci: *Ssa197*, *Ssa202*, *SSsp1605*, *SSsp2201*, *SSsp2213*, *SSsp2215*, *SSsp2216* and *SSspG7* (O'Reilly *et al.* 1996; Paterson *et al.* 2004). Detailed information about PCR protocols are described in Bobrowski (2010). Briefly, the eight microsatellites were amplified in 5 PCR reactions using fluorescent-labeled primers. The PCR products were run on an AB3730 DNA Analyzer (Applied Biosystems, USA) and the genotypes were analyzed using GeneMapper version 3.1

(Applied Biosystems, USA).  $F_{ST}$  and its 99% confidence interval was estimated using Fstat version 2.9.3.2 (Goudet 1995). The sample size for the  $F_{ST}$  estimate (520 fish) is much larger than that of the  $P_{ST}$  estimate (31 fish). To account for potential bias caused by sample size, I randomly selected microsatellite genotypes for 16 fish from each population (n=32 fish total) to calculate  $F_{ST}$  and replicated this analysis 1 000 times using the *pegas* package of R (Paradis 2010).

#### *P<sub>ST</sub> estimation*

$P_{ST}$ , the phenotypic analogue of  $Q_{ST}$ , is a measurement of phenotypic differentiation among populations. To calculate  $P_{ST}$  for each gene, I used the following model to obtain variance estimates between and within populations using restricted maximum likelihoods (REML) as priors:

$$Y_{ijklm} = \mu + P_i + T_j + Ba_k + Bl_l + e_{ijklm} \quad (3)$$

where  $Y_{ijklm}$  is the  $\log_2$  transformed normalized intensity value for each spot;  $\mu$  is the average value;  $P_i$  is the  $i^{\text{th}}$  effect of population;  $T_j$  is the  $j^{\text{th}}$  effect of tank;  $Ba_k$  is the  $k^{\text{th}}$  effect of batch;  $Bl_l$  is the  $l^{\text{th}}$  effect of block and  $e_{ijklm}$  is the random residual. I then used the variance estimates to calculate highest probability density (HPD) values with Markov Chain Monte Carlo (MCMC) simulations (10 000 replications) in the *languageR* R package (Baayen 2008). The median HPD values were used to calculate  $P_{ST}$  as:

$$P_{ST} = \sigma_{GB}^2 / (\sigma_{GB}^2 + 2\sigma_{GW}^2) \quad (4)$$

where  $\sigma_{GB}^2$  is median HPD value for the between-population variance and  $\sigma_{GW}^2$  is median HPD value for within-population variance.

## RESULTS

### *Tank effect on gene transcription*

In total, 271 genes were analyzed as these genes had expression data in more than 70% of the spots. Nineteen genes showed significant differences between the two replicate artificial stream tanks for the Sebago population (Table 3.1; Supplemental Figure S3.1). Fifteen genes showed significant differences between the two replicate stream tanks for the LaHave population (Table 3.2; Supplemental Figure S3.2). Among the identified genes, *proprotein convertase subtilisin/kexin type 5 (pcsk5)* and *tissue metalloproteinase inhibitor 3 precursor (timp3)* showed stream tank effects in both populations.

### *Differentially expressed genes between populations*

Twenty-one genes showed significantly different transcription between LaHave and Sebago populations, which accounts for 7.75% of the genes examined (Table 3.3; Supplemental Figure S3.3). Of the genes that were differentially transcribed between populations, 13 showed higher transcription levels in Sebago than that in LaHave whereas the other eight genes showed higher transcription levels in LaHave. Of these differentially expressed genes, *glutamine synthetase (glns)* and *myosin light chain 6B (myl6b)* were also affected by tank effect. Hierarchical clustering based on the 21 differentially expressed genes showed that the two replicate tanks within each population clustered together (Figure 3.1). The differentially expressed genes have diverse functions: nine genes (*cyp3a27*, *cyp2f5*, *fah*, *glns*, *hmox*, *hyal2*, *pded*, *pgm2* and *srk2tk*) encode enzymes; five genes (*clqc*, *il1r2*, *saa5*, *tcrb* and *tnr5*) are involved in the immune

response; five genes (*myl1*, *myl3*, *myl6b*, *fah* and *pgm2*) are involved in ion binding; two genes (*cytl1* and *tbl1xr1b*) regulate transcription activity and one gene (*grn*) regulates cell growth.

#### *F<sub>ST</sub>* and *P<sub>ST</sub>*

The *F<sub>ST</sub>* value based on the microsatellite genotypes of all 520 fish was 0.038, with a 99% confidence interval of 0.020-0.057. The mean *F<sub>ST</sub>* value ( $\pm$ SD) based on the randomly sub-sampled microsatellite genotypes of 32 fish was  $0.037 \pm 0.008$ . Although the *F<sub>ST</sub>* estimates based on the randomly sub-sampled fish ranged from 0.013 to 0.082, 97.5% (975 out of 1000 times) of the estimates were within the 99% confidence interval of the mean *F<sub>ST</sub>* value based on all 520 fish. This result indicates that *F<sub>ST</sub>*-*P<sub>ST</sub>* comparison is not likely biased due to estimates based on different sample size.

I tested for the effects of selection versus genetic drift as contributors to the difference in gene transcription levels by comparing *P<sub>ST</sub>* for each gene with the *F<sub>ST</sub>* confidence interval. The *P<sub>ST</sub>* for the 271 analyzed genes ranged from 0.034 to 0.32 (Figure 3.2a). The *P<sub>ST</sub>* for the 21 differentially expressed genes between populations ranged from 0.20 to 0.32 (Figure 3.2b), all of which were substantially outside the 99% confidence interval for the *F<sub>ST</sub>* value reported above. I therefore conclude that the population difference in transcription for these 21 genes is primarily driven by selection.

## DISCUSSION

Gene transcription data has only recently been used in the study of population genetics. Comparing gene transcription profiles across populations does have important

applications in conservation and management (Vandersteen Tymchuk *et al.* 2010). In this experiment, I compared transcription between two Atlantic salmon populations (LaHave and Sebago), and found that about 8% of analyzed genes were differentially expressed between the two populations, despite being held in identical artificial environments. This percentage is higher than that reported in two similar studies (1.4% and 1.7%: Roberge *et al.* 2006; and 2.3%: Debes *et al.* 2012) on microarray gene transcription comparisons between farmed and wild Atlantic salmon. There are three possibilities for my higher frequency of transcriptional differences. First, my custom microarray was enriched for genes that are known to be sensitive to environmental differences. Second, the two populations I compared have different evolutionary histories and marked life history differences. Third, I used an oligonucleotide microarray which may be more sensitive than the cDNA microarrays used in other studies (Yauk *et al.* 2004).

Like other quantitative traits, gene expression is determined by a combination of genetic and environmental effects, thus it is not surprising that I detected both tank and population effects. Two genes, *psck5* and *timp3*, were significantly affected by tank effects in both populations, while the majority of genes differentially expressed between tanks showed difference in only one population. Thus minor environmental differences among tanks affected the two populations differently, likely a reflection of genotype by environmental interactions (G×E) on gene transcription. This is despite my attempts to control many environmental factors, for example: the fish were crossed on the same day and reared under the same food and water source and flow regimes, plus I sampled them at the same developmental stage and used identical protocols to measure gene transcription. Nevertheless, stream tank effects contributed to differences in gene



expression, and the number of genes and magnitude of differences between tanks was similar to the population effect. As my design had the stream tanks nested within populations, I am unable to specifically partition G×E effects, however, previous studies have shown that G×E contributes to transcriptional variation (Smith & Kruglyak 2008; Grishkevich & Yanai 2013). Although I cannot definitively conclude that the stream tank effects reflected G×E, the transcription differences indicate high environmental sensitivity in these fish, perhaps reflecting why reintroduction may succeed in one habitat but fail in another using the same donor stock.

Neutral microsatellite DNA markers have been widely used in conservation genetics over the past two decades under the assumption that the extent of neutral genetic variation is positively correlated with the genome-wide functional genetic variation - this assumption has been called into question in a number of studies (Reed & Frankham 2001; Hedrick 2001). In this study, I found that there was little neutral genetic differentiation ( $F_{ST} = 0.038$ ) between the two populations based on microsatellite genotypes, but genes involved in known and vital functions showed significant differences between the two populations. For example, two genes (*cyp3a27* and *cyp2f5*) encoding cytochrome P450 (CYP) enzymes had higher transcription in the anadromous population (LaHave); those enzymes play an important role in metabolism of steroids and fatty acids and detoxification of pollutants and drugs (Uno *et al.* 2012). Similarly, *glns* which encodes glutamine synthetase (catalyzes ammonia and glutamate to synthesize glutamine) also showed higher transcription levels in the LaHave population. The conversion of ammonia to glutamine is a mechanism to remove ammonia and thus avoid its toxicity (Essex-Fraser *et al.* 2005). The higher transcription of the CYP and *glns* genes may be adaptive

for the anadromous population as part of their preparation for the novel marine environment. Similar migratory preparation was also reported in Giger *et al.* (2008) where 17 genes related to migratory adaptation differentially expressed between migratory and non-migratory brown trout populations. In contrast, I found 5 immune-related genes (*tnfr5*, *tcrb*, *il1r2*, *clqc* and *saa5*) had higher transcription levels in the Sebago relative to the LaHave population. The *tnfr5* gene encodes a member of the tumor necrosis factor receptor superfamily and the interaction between the receptor and its ligand plays a crucial role in expression regulation of many immune molecules, such as cytokines and chemokines (Chatzigeorgiou *et al.* 2009). The *tcrb* gene encodes the  $\beta$  chain of the T cell receptor in  $\alpha\beta$  T cells which recognizes foreign antigens that are bound by major histocompatibility complex molecules (Goldrath & Bevan 1999). The *il1r2* gene encodes interleukin 1 receptor 2 which binds and inhibits interleukin 1 activity (Colotta *et al.* 1993). The *clqc* gene encodes the C-chain of complement subcomponent C1q. C1q is the recognition subunit of the C1 complex and is able to recognize and bind a variety of targets to activate the complement pathway to defense pathogens (Gaboriaud *et al.* 2004). The *saa5* gene encodes an acute phase protein which is involved in the inflammatory response and lipid transportation (Banka *et al.* 1995; Goltry *et al.* 1998), and this gene is known to be up-regulated after bacterial and viral infection (Miwata *et al.* 1993; Lin *et al.* 2007). The different transcription of those immune genes is related to coping with pathogens in their environments, which is vital for salmonid survival in the wild (Miller *et al.* 2011). Previous reintroduction of Atlantic salmon into Lake Ontario focused on LaHave population which has already been identified as a possibly inappropriate population (Van Zwol *et al.* 2012), and as my data show, the LaHave

population, while showing higher expression of CYP genes, appears to have lower expression at selected immune genes. Thus my transcriptional profiling of functionally important genes shows that not only is gene expression variation more divergent between the populations than expected based on drift (neutral DNA), but that the Sebago Atlantic salmon population may be a better choice for reintroduction into Lake Ontario.

Identifying genes with transcription profiles that indicate selection-based differences among populations is important in conservation and management as such differences likely underlie adaptations to different environmental conditions. In my study, the  $P_{ST}$  values calculated were comparable to  $P_{ST}$  and  $Q_{ST}$  values for transcription in rainbow trout (Aykanat *et al.* 2011; Wellband & Heath 2013), but much higher than the  $Q_{ST}$  values estimated in two Atlantic salmon subpopulations (Roberge *et al.* 2007), which implies that differences in gene transcription among populations depend on the extent of divergence. My results showed that much of the difference in gene transcription between the two populations of Atlantic salmon was consistent with divergence by selection. Moreover, the genes identified as driven by directional selection are excellent candidate markers for predicting fitness in specific environments.

Although the application of gene transcription in conservation biology is still in its infancy, transcriptional profiling of potential source populations can enhance reintroduction efforts in two ways: first, gene expression comparisons can identify functional differences that are related to important physiological processes and responses to environmental stressors, and subsequently, variation in individual gene transcription can be used to predict specific trait response upon reintroduction (Miller *et al.* 2011). The custom DNA microarray I developed provides a relatively inexpensive method to profile

transcription for many individuals that will make it possible to choose appropriate source population for reintroduction. Such an approach will increase the likelihood of reintroduction success and ultimately, conservation. Furthermore, as more such studies are completed, and our understanding of the role of specific gene expression responses in adaptive environmental stress responses improves, the application of transcriptional profiling will expand.

## REFERENCES

- Aykanat T, Thrower FP, Heath DD (2011) Rapid evolution of osmoregulatory function by modification of gene transcription in steelhead trout. *Genetica*, **139**, 233–242.
- Baayen RH (2008) *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge University Press, Cambridge.
- Banka CL, Yuan T, de Beer MC *et al.* (1995) Serum amyloid A (SAA): influence on HDL-mediated cellular cholesterol efflux. *Journal of Lipid Research*, **36**, 1058–1065.
- Bates D, Maechler M (2010) *lme4: Linear mixed-effects models using Eigen and Eigen*. R package version 0.999375-32. <https://cran.r-project.org/package=lme4>.
- Bobrowski R (2009) Survival, growth, and out-migration timing of reintroduced Atlantic salmon (*Salmo salar*) in Cobourg Brook, Ontario. M.Sc. Thesis. Trent University, Canada.
- Campos-Perez JJ, Ward M, Grabowski PS, Ellis AE, Secombes CJ (2000) The gills are an important site of iNOS expression in rainbow trout *Oncorhynchus mykiss* after challenge with the gram-positive pathogen *Renibacterium salmoninarum*. *Immunology*, **99**, 153–161.
- Chatzigeorgiou A, Lyberi M, Chatzilymperis G, Nezos A, Kamper E (2009) CD40/CD40L signaling and its implication in health and disease. *BioFactors*, **35**, 474–483.
- Coghlan SM, Ringler NH (2004) A comparison of Atlantic salmon embryo and fry stocking in the Salmon River, New York. *North American Journal of Fisheries Management*, **24**, 1385–1397.
- Colotta F, Re F, Muzio M *et al.* (1993) Interleukin-1 type II receptor: a decoy target for IL-1 that is regulated by IL-4. *Science*, **261**, 472–475.
- Crawford SS (2001) Salmonine introductions to the Laurentian Great Lakes: an historical review and evaluation of ecological effects. *Canadian Special Publication of Fisheries and Aquatic Sciences* 132, 205 pp.
- Debes PV, Normandeau E, Fraser DJ, Bernatchez L, Hutchings JA (2012) Differences in transcription levels among wild, domesticated, and hybrid Atlantic salmon (*Salmo salar*) from two environments. *Molecular Ecology*, **21**, 2574–2587.

Dimond P, Smitka J (2005) Evaluation of selected strains of Atlantic salmon as potential candidates for the restoration of Lake Ontario. *Trout Unlimited Canada Technical Report* ON-012, 41 pp.

Essex-Fraser PA, Steele SL, Bernier NJ *et al.* (2005) Expression of four glutamine synthetase genes in the early stages of development of rainbow trout (*Oncorhynchus mykiss*) in relationship to nitrogen excretion. *The Journal of Biological Chemistry*, **280**, 20268–20273.

Evans DH, Piermarini PM, Choe KP (2005) The multifunctional fish gill: dominant site of gas exchange, osmoregulation, acid-base regulation, and excretion of nitrogenous waste. *Physiological Reviews*, **85**, 97–177.

Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies*, **326 Suppl 1**, S22–29.

Frankham R, Briscoe DA, Ballou JD (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge.

Gaboriaud C, Thielens NM, Gregory LA *et al.* (2004) Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends in Immunology*, **25**, 368–373.

Garcia de Leaniz C, Fleming IA, Einum S *et al.* (2007) A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. *Biological Reviews of the Cambridge Philosophical Society*, **82**, 173–211.

Giger T, Excoffier L, Amstutz U *et al.* (2008) Population transcriptomics of life-history variation in the genus *Salmo*. *Molecular Ecology*, **17**, 3095–3108.

Giger T, Excoffier L, Day PJR *et al.* (2006) Life history shapes gene expression in salmonids. *Current Biology*, **16**, R281–282.

Goldrath AW, Bevan MJ (1999) Selecting and maintaining a diverse T-cell repertoire. *Nature*, **402**, 255–262.

Goltry KL, Epperly MW, Greenberger JS (1998) Induction of serum amyloid A inflammatory response genes in irradiated bone marrow cells. *Radiation Research*, **149**, 570–578.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Grishkevich V, Yanai I (2013) The genomic determinants of genotype  $\times$  environment interactions in gene expression. *Trends in Genetics*, **29**, 479–487.

- Hansen MM (2010) Expression of interest: transcriptomics and the designation of conservation units. *Molecular Ecology*, **19**, 1757–1759.
- Hedrick PW (2001) Conservation genetics: Where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.
- Houde AL, Wilson CC, Neff BD (2013) Genetic architecture of survival and fitness-related traits in two populations of Atlantic salmon. *Heredity*, **111**, 513–519.
- Houde ALS, Wilson CC, Neff BD (2015) Competitive interactions among multiple non-native salmonids and two populations of Atlantic salmon. *Ecology of Freshwater Fish*, **24**, 44–55.
- Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, **4**, 44–57.
- Hutter S, Saminadin-Peter SS, Stephan W, Parsch J (2008) Gene expression variation in African and European populations of *Drosophila melanogaster*. *Genome Biology*, **9**, R12.
- Lamaze FC, Garant D, Bernatchez L (2013) Stocking impacts the expression of candidate genes and physiological condition in introgressed brook charr (*Salvelinus fontinalis*) populations. *Evolutionary Applications*, **6**, 393–407.
- Larsen PF, Nielsen EE, Williams TD *et al.* (2007) Adaptive differences in gene expression in European flounder (*Platichthys flesus*). *Molecular Ecology*, **16**, 4674–4683.
- Lin B, Chen S, Cao Z *et al.* (2007) Acute phase response in zebrafish upon *Aeromonas salmonicida* and *Staphylococcus aureus* infection: striking similarities and obvious differences with mammals. *Molecular Immunology*, **44**, 295–301.
- Luca F, Kashyap S, Southard C *et al.* (2009) Adaptive variation regulates the expression of the human *SGK1* gene in response to stress. *PLoS Genetics*, **5**, e1000489.
- Miller KM, Li S, Kaukinen KH *et al.* (2011) Genomic signatures predict migration and spawning failure in wild Canadian salmon. *Science*, **331**, 214–217.
- Miwata H, Yamada T, Okada M *et al.* (1993) Serum amyloid A protein in acute viral infections. *Archives of Disease in Childhood*, **68**, 210–214.
- Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.
- O'Reilly PT, Hamilton LC, McConnell SK, Wright JM (1996) Rapid analysis of genetic variation in Atlantic salmon (*Salmo salar*) by PCR multiplexing of dinucleotide and

tetranucleotide microsatellites. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 2292–2298.

Ouborg NJ, Pertoldi C, Loeschcke V, Bijlsma RK, Hedrick PW (2010) Conservation genetics in transition to conservation genomics. *Trends in Genetics*, **26**, 177–187.

Paradis E (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, **26**, 419–420.

Parrish DL, Behnke RJ, Gephard SR, McCormick SD, Reeves GH (1998) Why aren't there more Atlantic salmon (*Salmo salar*)? *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 281–287.

Paterson S, Piernney SB, Knox D, Gilbey J, Verspoor E (2004) Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. *Molecular Ecology Notes*, **4**, 160–162.

Pedersen KS, Kristensen TN, Loeschcke V (2005) Effects of inbreeding and rate of inbreeding in *Drosophila melanogaster*- Hsp70 expression and fitness. *Journal of Evolutionary Biology*, **18**, 756–762.

Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution*, **55**, 1095–1103.

Roberge C, Einum S, Guderley H, Bernatchez L (2006) Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. *Molecular Ecology*, **15**, 9–20.

Roberge C, Guderley H, Bernatchez L (2007) Genomewide identification of genes under directional selection: gene transcription  $Q_{ST}$  scan in diverging Atlantic salmon subpopulations. *Genetics*, **177**, 1011–1022.

Roberge C, Normandeau E, Einum S, Guderley H, Bernatchez L (2008) Genetic consequences of interbreeding between farmed and wild Atlantic salmon: insights from the transcriptome. *Molecular Ecology*, **17**, 314–324.

Rouillard J-M, Zuker M, Gulari E (2003) OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Research*, **31**, 3057–3062.

Saeed AI, Sharov V, White J *et al.* (2003) TM4: a free, open-source system for microarray data management and analysis. *BioTechniques*, **34**, 374–378.

Schlichting CD, Smith H (2002) Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, **16**, 189–211.



- Scott RJ, Poos MS, Noakes DLG, Beamish FWH (2005) Effects of exotic salmonids on juvenile Atlantic salmon behaviour. *Ecology of Freshwater Fish*, **14**, 283–288.
- Smith EN, Kruglyak L (2008) Gene–environment interaction in yeast gene expression. *PLoS Biology*, **6**, e83.
- Smyth GK (2005) limma: Linear Models for Microarray Data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* Statistics for Biology and Health. (eds Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S), pp. 397–420. Springer, New York.
- Spielman D, Brook BW, Briscoe DA, Frankham R (2004) Does inbreeding and loss of genetic diversity decrease disease resistance? *Conservation Genetics*, **5**, 439–448.
- Storey JD, Madeoy J, Strout JL *et al.* (2007) Gene-expression variation within and among human populations. *American Journal of Human Genetics*, **80**, 502–509.
- Townsend JP, Cavalieri D, Hartl DL (2003) Population genetic variation in genome-wide gene expression. *Molecular Biology and Evolution*, **20**, 955–963.
- Uno T, Ishizuka M, Itakura T (2012) Cytochrome P450 (CYP) in fish. *Environmental Toxicology and Pharmacology*, **34**, 1–13.
- Van Zwol JA, Neff BD, Wilson CC (2012) The influence of non-native salmonids on circulating hormone concentrations in juvenile Atlantic salmon. *Animal Behaviour*, **83**, 119–129.
- Vandersteen Tymchuk W, O'Reilly P, Bittman J, Macdonald D, Schulte P (2010) Conservation genomics of Atlantic salmon: variation in gene expression between and within regions of the Bay of Fundy. *Molecular Ecology*, **19**, 1842–1859.
- Wellband KW, Heath DD (2013) The relative contribution of drift and selection to transcriptional divergence among Babine Lake tributary populations of juvenile rainbow trout. *Journal of Evolutionary Biology*, **26**, 2497–2508.
- Whitehead A, Crawford DL (2006) Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 5425–5430.
- Yauk CL, Berndt ML, Williams A, Douglas GR (2004) Comprehensive comparison of six microarray technologies. *Nucleic Acids Research*, **32**, e124.

**Table 3.1** List of genes showing significantly different transcription between the two replicate tanks for the Sebago population of Atlantic salmon (*Salmo salar*).

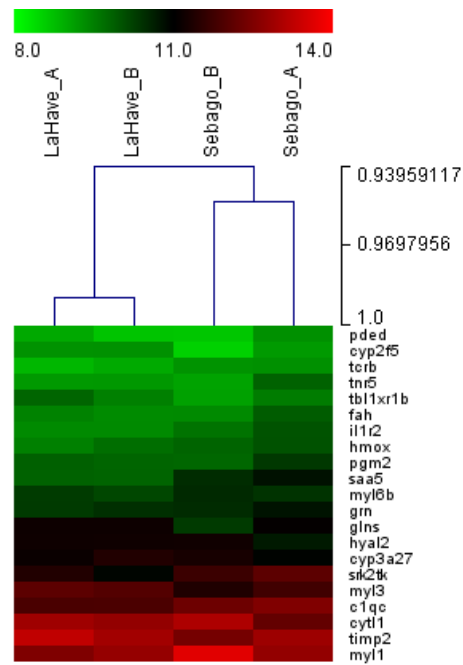
Gene symbol	Gene product	Intensity ratio (Tank1/Tank2)
<i>cish</i>	cytokine-inducible SH2-containing protein	0.51
<i>cfb</i>	complement factor B precursor	0.60
<i>pcsk5</i>	proprotein convertase subtilisin/kexin type 5 precursor	0.67
<i>pitpna</i>	phosphatidylinositol transfer protein alpha isoform	0.74
<i>glns</i>	glutamine synthetase	0.76
<i>timp3</i>	metalloproteinase inhibitor 3 precursor	0.76
<i>atl2</i>	atlastin-2	0.77
<i>psmd5</i>	26S proteasome non-ATPase regulatory subunit 5	0.77
<i>gem</i>	GTP-binding protein	1.25
<i>xaf1</i>	XIAP-associated factor 1	1.27
<i>psmc2</i>	26S protease regulatory subunit 7	1.30
<i>mstn</i>	myostatin 1b	1.33
<i>crtam</i>	cytotoxic and regulatory T-cell molecule precursor	1.35
<i>hspa14</i>	heat shock 70 kDa protein 14	1.37
<i>sod3</i>	extracellular superoxide dismutase [Cu-Zn] precursor	1.46
<i>hsf2</i>	heat shock factor protein 2	1.51
<i>sar1a</i>	GTP-binding protein SAR1a	1.53
<i>cdk5</i>	cell division protein kinase 5	1.60
<i>pcna</i>	proliferating cell nuclear antigen putative mRNA	1.61

**Table 3.2** List of genes showing significantly different transcription between the two replicate tanks for the LaHave population of Atlantic salmon (*Salmo salar*).

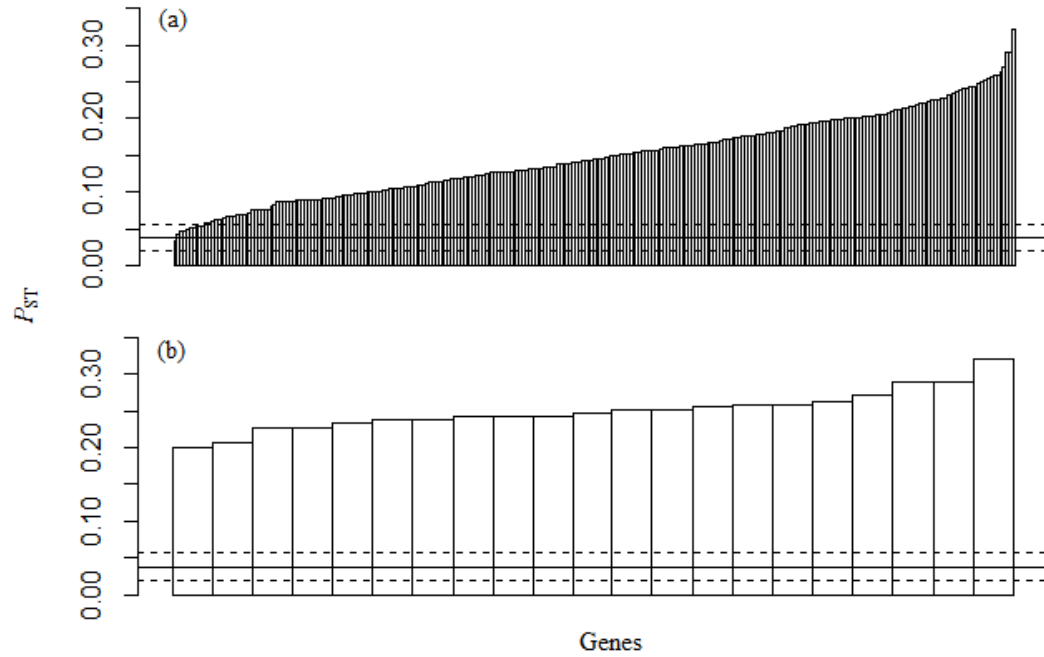
Gene symbol	Gene product	Intensity ratio (Tank3/Tank4)
<i>sar1b</i>	GTP-binding protein SAR1b	0.71
<i>pcsk5</i>	proprotein convertase subtilisin/kexin type 5 precursor	0.79
<i>myl6b</i>	myosin light chain 6B	0.88
<i>irak3</i>	interleukin-1 receptor-associated kinase 3	1.15
<i>psmd9</i>	26S proteasome non-ATPase regulatory subunit 9	1.20
<i>isca2</i>	iron-sulfur cluster assembly 2 homolog, mitochondrial precursor	1.26
<i>atp1a1</i>	sodium/potassium-transporting ATPase subunit alpha-1 precursor	1.26
<i>tmp49</i>	transmembrane protein 49	1.27
<i>ptgd2</i>	glutathione-requiring prostaglandin D synthase	1.27
<i>cdk9</i>	cell division protein kinase 9	1.30
<i>timp3</i>	metalloproteinase inhibitor 3 precursor	1.31
<i>il4</i>	interleukin 4/13A (il4/13a)	1.42
<i>ctsh</i>	cathepsin H precursor	1.42
<i>pgd</i>	6-phosphogluconate dehydrogenase, decarboxylating	1.54
<i>c7</i>	complement C7 precursor	1.54

**Table 3.3** List of genes showing significantly different transcription between the Sebago and LaHave populations of Atlantic salmon (*Salmo salar*).

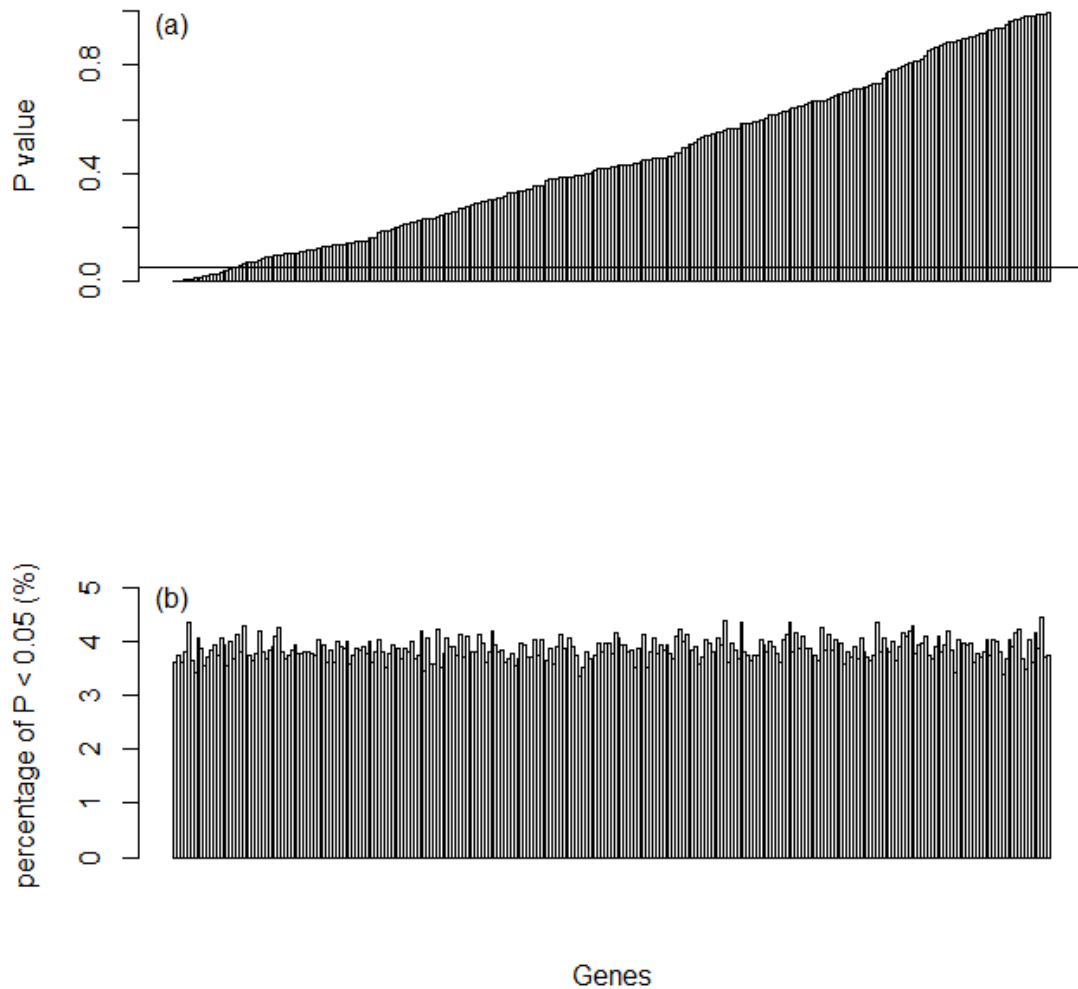
Gene symbol	Gene product	Intensity ratio (Sebago/LaHave)
<i>timp2</i>	TIMP Metalloproteinase inhibitor 2 precursor	0.72
<i>myl3</i>	myosin light chain 3	0.76
<i>glns</i>	glutamine synthetase	0.76
<i>tbl1xr1b</i>	F-box-like/WD repeat-containing protein TBL1XR1-B	0.77
<i>cyp3a27</i>	cytochrome P450 3A27	0.79
<i>cytl1</i>	cytokine-like protein 1 precursor	0.79
<i>cyp2f5</i>	cytochrome P450 2F5	0.79
<i>hyal2</i>	hyaluronidase-2	0.81
<i>pded</i>	phosphodiesterase delta-like protein	1.18
<i>pgm2</i>	phosphoglucomutase-2	1.20
<i>grn</i>	granulins precursor	1.21
<i>myl6b</i>	myosin light chain 6B	1.22
<i>fah</i>	fumarylacetoacetate hydrolase	1.23
<i>tnr5</i>	tumor necrosis factor receptor superfamily member 5	1.25
<i>hmox</i>	heme oxygenase	1.26
<i>tcrb</i>	T-cell receptor beta chain	1.28
<i>myl1</i>	myosin light chain 1, skeletal muscle isoform	1.33
<i>il1r2</i>	interleukin-1 receptor type II precursor	1.44
<i>c1qc</i>	complement C1q subcomponent subunit C precursor	1.47
<i>srk2tk</i>	SRK2 tyrosine kinase	1.70
<i>saa5</i>	serum amyloid A-5 protein	1.83



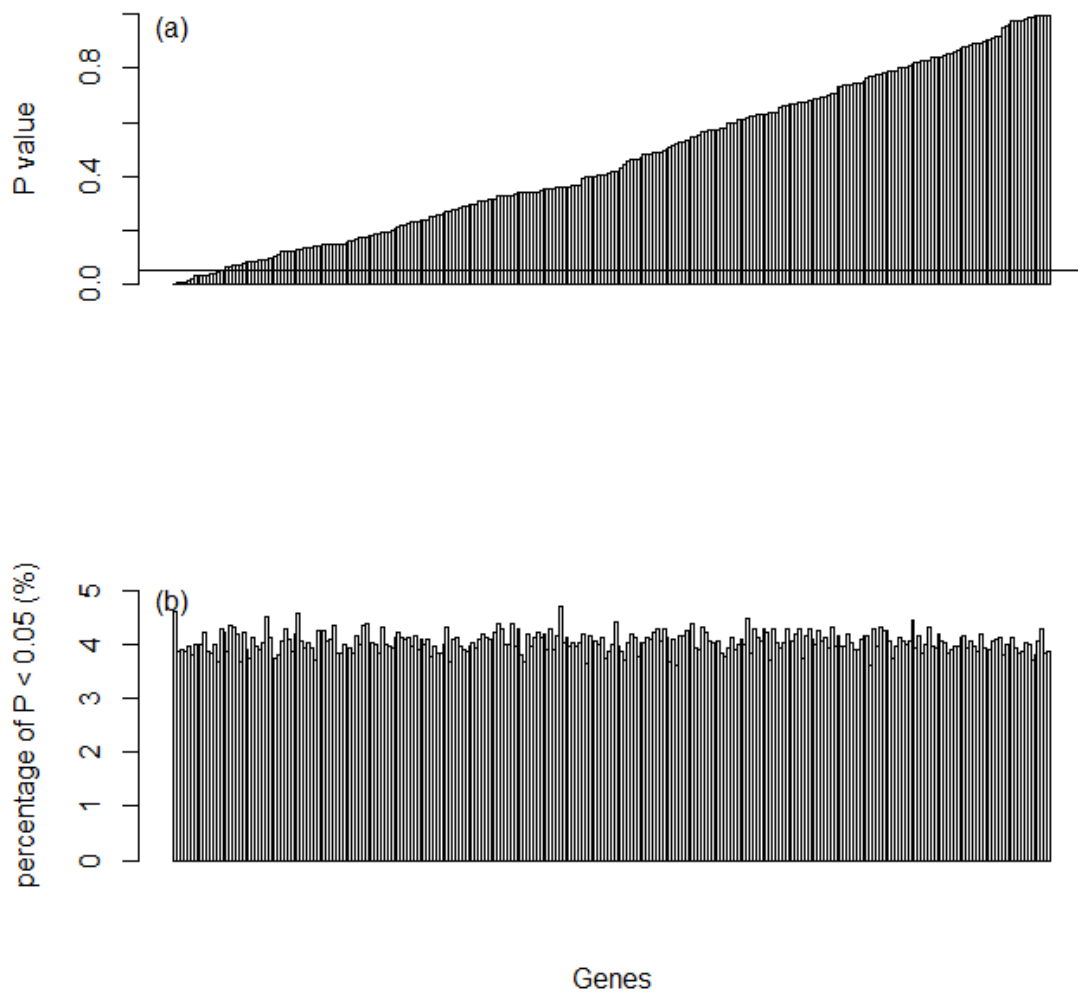
**Figure 3.1** Gene transcription heatmap showing hierarchical clustering of the 21 differentially expressed genes between Sebago and LaHave Atlantic salmon (*Salmo salar*) populations. The rows represent different genes and the columns represent different tanks. The transcription level for each gene is the average  $\log_2$  transformed intensity value of fish from the same tank.



**Figure 3.2** Histograms of global  $P_{ST}$  for transcription of genes between two populations of juvenile Atlantic salmon (*Salmo salar*) arranged in increasing order. Panel a:  $P_{ST}$  values for all 271 analyzed genes. Panel b:  $P_{ST}$  values for genes which showed significantly different transcription levels between the two populations. The two horizontal dashed lines represent the upper and lower limits of the 99% confidence interval of  $F_{ST}$  based on microsatellite genotypes at eight loci. The solid line represents the mean  $F_{ST}$ .

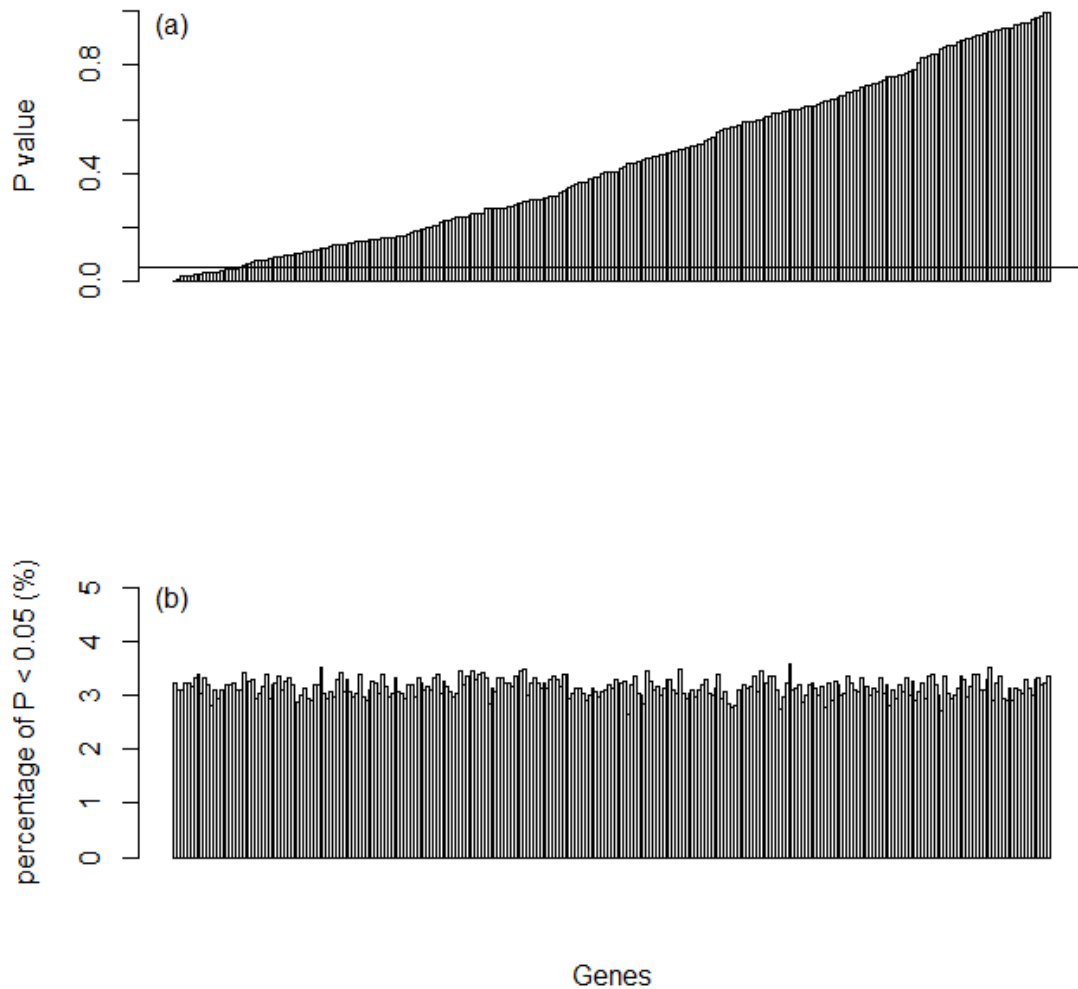


**Supplementary Figure S3.1** Histograms of  $P$  values for the analysis of difference in gene transcription between two tanks of the Sebago Atlantic salmon (*Salmo salar*) population. Panel a:  $P$  value for each gene arranged in increasing order. Panel b: percentage of  $P$  value for each gene with the same order as (a) that is less than 0.05 during the 10 000 permutation. The solid line in (a) represents the statistical significance (0.05).



**Supplementary Figure S3.2** Histograms of  $P$  values for the analysis of difference in gene transcription between two tanks of the LaHave Atlantic salmon (*Salmo salar*) population. Panel a:  $P$  value for each gene arranged in increasing order. Panel b: percentage of  $P$  value for each gene with the same order as (a) that is less than 0.05 during the 10 000 permutation. The solid line in (a) represents the statistical significance (0.05).





**Supplementary Figure S3.3** Histograms of  $P$  values for the analysis of difference in gene transcription between Sebago and LaHave Atlantic salmon (*Salmo salar*) populations. Panel a:  $P$  value for each gene arranged in increasing order. Panel b: percentage of  $P$  value for each gene with the same order as (a) that is less than 0.05 during the 10 000 permutation. The solid line in (a) represents the statistical significance (0.05).

## CHAPTER 4

### GENETIC ARCHITECTURE OF GENE TRANSCRIPTION IN TWO ATLANTIC SALMON POPULATIONS<sup>3</sup>

#### INTRODUCTION

Acclimation via phenotypic plasticity and adaptation via selection are the two main mechanisms organisms use to cope with environmental stress, and they thus play a vital role in organism survival and population persistence in changing or novel environments (Hansen *et al.* 2012). Gene expression regulation has long been recognized as playing an important role in acclimation and adaptation (Hochachka & Somero 1984; Crawford & Powers 1992). The ability of individual genotypes to express different phenotypes in response to environmental fluctuations and stress challenges is mediated by regulation in gene expression, specifically through quantitative changes in gene expression levels and selective expression of different isoforms (Schlichting & Smith 2002; Schulte 2004). While gene expression is influenced by internal and external environmental factors (López-Maury *et al.* 2008; Hodgins-Davis & Townsend 2009), genetic analysis has shown that gene expression variation is heritable and gene expression levels are also determined by genetic factors (Cheung and Spielman, 2002; Buckland, 2004; Petretto *et al.*, 2006). Moreover, it was shown over four decades ago that evolutionary differences among species may depend more on changes in gene expression regulation than changes in gene sequences (King & Wilson 1975). Recently,

---

<sup>3</sup> This is the outcome of joint research.

novel genomic tools have shown that local adaptation among populations within a species are ten-times more likely to be driven by genetic variation regulating gene expression than that by genetic variation changing amino acid sequences (Fraser 2013). Therefore, gene expression is not only a process whereby organisms acclimate to new or changing environments, but also provides a source of variation that selection can act on.

Populations of the same species can have different responses and tolerances to environmental stressors due to their separate evolutionary histories, resulting in divergent performance that is relevant for both culture and conservation in the field (Forsman 2014; He *et al.* 2016). Population-level differences in tolerance of environmental toxins and temperature stress has been reported to be associated with differences in regulation and pre-adaptation of gene expression, reflecting local habitats conditions (Whitehead *et al.* 2010; Schoville *et al.* 2012; Gleason & Burton 2015). Although there are many studies that report population differences in gene expression, comparisons of genetic components of gene expression variance among populations have received much less attention.

Additive genetic, non-additive genetic, and maternal effects are important for evolutionary processes due to their contributions to phenotypic variation. Additive genetic effects are critical for evolutionary responses and are directly related to evolutionary rate in response to selection according to the breeder's equation (Falconer and Mackay 1996). Non-additive genetic effects (e.g., dominance and epistasis) can significantly contribute to phenotypic variation (Evans & Neff 2009; Gallardo *et al.* 2010); and non-additive effects increase in extreme environments and in response to stress (Jinks *et al.* 1973; Aykanat *et al.* 2012b). Maternal effects are significant contributors to variation in fitness-related traits at early life stages (Heath *et al.* 1999;

Aykanat *et al.* 2012a; Houde *et al.* 2013). Populations of the same species subject to different evolutionary histories and selection pressures can have different genetic architectures for fitness-related traits. Although gene expression is closely linked to phenotype and can be highly heritable, it is not clear how populations differ in additive genetic components of gene expression variation. Similarly, although maternal effects and non-additive genetic effects play important roles in individual and population performance, the contribution of those effects to gene expression variation among populations is unknown.

The aim of this study is to investigate the genetic architecture of gene transcription and to test whether and how the genetic architecture for key muscle function-related genes varies between populations with different life histories. To this end, I implemented a 5×5 full factorial breeding design for each of two Atlantic salmon (*Salmo salar*) populations: LaHave (anadromous) and Sebago (landlocked). Atlantic salmon have important roles in aquaculture, and the commercial and recreational fisheries; however, they are of conservation concern across most of their native range (Parrish *et al.* 1998). I used a high throughput qRT-PCR system to quantify gene transcription of 22 genes in muscle for over 400 juvenile Atlantic salmon. I analyzed population and parental effects on gene transcription and partitioned variance into additive genetic, non-additive genetic, and maternal effects within each population.

## MATERIALS AND METHODS

### *Breeding design*

I used parental fish from two Atlantic salmon populations (LaHave and Sebago) which were provided by the Ontario Ministry of Natural Resources and Forestry (OMNRF). The LaHave population originates from the LaHave River, Nova Scotia; fertilized eggs from wild LaHave Atlantic salmon population were received by OMNRF in 1995 (Houde *et al.* 2015). The Sebago population originates from Sebago Lake, Maine; fertilized eggs from a hatchery supplemented wild Sebago Atlantic salmon population were received by OMNRF in 2006 (Houde *et al.* 2015). Thus the LaHave was fourth generation hatchery and Sebago was second generation hatchery within the OMNRF facility at the time of this study. The  $F_{ST}$  between these two populations is 0.038 (He *et al.* 2015), and a population genetic study showed only small genetic diversity (heterozygosity) differences between anadromous and landlocked Atlantic salmon populations in North America (King *et al.* 2001). In early November 2011, I used 5 males and 5 females from each population where each male was crossed to each female in a full-factorial design to produce 25 families. The fertilized eggs were incubated in vertical stack incubators in replicate cells. After hatching, each family was reared in two replicate rearing tanks. In May 2012 (187 days post fertilization), three to five fry per tank were humanely euthanized by overdose of MS222 and the whole fish was preserved in RNA later. All 50 families had representative fish from both replicate tanks except one family in the LaHave population which had fish sampled from only one tank.

### *RNA extraction and cDNA preparation*

Although this study was designed to target muscle tissue, the fish were too small (average body mass:  $0.34 \pm 0.11$ g) to collect pure muscle tissue from preserved samples. Thus to collect tissue samples for RNA isolation, I first cut the fish through the lateral line and then used the tissue between the dorsal and adipose fins for RNA extraction. Therefore, the tissue I used for RNA extraction in this study was primarily muscle, however the sample also contained skin and some bone tissue. The tissue sample was mechanically homogenized in the presence of Isol-RNA Lysis Reagent (5 PRIME, Gaithersburg, MD, USA) following the manufacturer's protocol. The concentration and purity of extracted RNA were checked using a NanoVue spectrophotometer (GE Healthcare Bio-Science Corp, USA). The quality of a subset of RNA samples was also checked using the Agilent RNA 6000 Nano Kit in an Agilent 2100 Bioanalyzer (Agilent Technologies, Mississauga, ON, Canada). About 2 µg RNA was treated with DNase I (Sigma-Aldrich, Oakville, ON, Canada) to remove contaminated DNA and then used for cDNA synthesis using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Burlington, ON, Canada). In total, I reverse transcribed 480 RNA samples. The quality of 192 cDNA samples was checked by PCR amplification of the *ef1ab* gene, and all showed the expected band on agarose gel.

### *Gene selection*

Fish muscle plays an important role in swimming performance and energy regulation, which is vital for fish to migrate, forage and avoid predation in the wild. Muscle accounts for 50 - 60% of the total body mass in salmonids (Jobling 1993), and

thus regulation of muscle development and growth is also very important for aquaculture. The 22 genes (Table 4.1) I chose for this study are involved in a variety of key processes in fish muscle. Seven genes encode enzymes directly involved in energy generation by participating in fatty acid oxidation (*acadl*, *acadm* and *cpt2*), citric acid cycle (*cs*), carbohydrate oxidation (*dlat*), anaerobic glycolysis (*ldha*), and glycogenolysis (*pygma*). Five genes are involved in energy regulation through signal transduction (*camk2g* and *prkag2*), fatty acid transport (*fabp3*) and transcription regulation (*ppara* and *tfam*). Six genes encode muscle structural proteins that are involved in muscle contraction (*dysi1*, *fmyhc*, *mlc2*, *smyhc1*, *tnni2* and *tnnt1*). The *dysi1* gene was selected in the muscle contraction gene group because it is involved in muscle membrane repair (Han & Campbell 2007). Four selected genes (*frap1*, *myf5*, *myod1* and *murf1*) are growth-related genes. The *myf5* and *myod1* genes are members of the myogenic regulatory factors and regulate muscle cell differentiation and growth (Francetic & Li 2011). The *murf1* gene belongs to the ubiquitin-protein ligases which are required for the ubiquitin proteasome system to degrade muscle proteins (Lecker 2003). I also included the *frap1* gene which has multiple cellular signal transduction roles in mediation of stress response and is known to regulate cell growth and proliferation (Desai *et al.* 2002; Murakami *et al.* 2004).

#### *Quantitative real-time PCR*

The sequences of TaqMan primers and probes for three genes were obtained from journal publications: *eflab* (Løvoll *et al.* 2011), *fabp3* (Torstensen *et al.* 2009) and *fmyhc* (Hevrøy *et al.* 2006). The sequences of the other genes were downloaded from NCBI

(<http://www.ncbi.nlm.nih.gov/>) and the primers and TaqMan MGB probes (Table 4.1) for those genes were designed using Primer Express 3.0. The primers are predicted to have 100% efficiency using the pcrEfficiency software (Mallona *et al.* 2011). The primers and probes were synthesized and spotted into through-holes of OpenArray chips by Applied Biosystems (Applied Biosystems, Burlington, ON, Canada).

OpenArray qRT-PCR was performed using a QuantStudio 12K Flex Real-Time PCR System following the manufacture's instruction. I prepared a 5  $\mu$ L mixture for each cDNA sample which contained 2.5  $\mu$ L TaqMan® OpenArray® Real-Time PCR Master Mix (Applied Biosystems, Burlington, ON, Canada) and 1.2  $\mu$ L cDNA. I used the 56×48 format OpenArray chip which has 48 subarrays in each chip and each subarray contains 64 through-holes. Each chip can be used to measure gene expression for 48 individual cDNA samples for all the targeted genes in duplicate. The 5  $\mu$ L mixtures were prepared in 384-well plates and were then loaded into OpenArray chips using the OpenArray AccuFill System and each qRT-PCR reaction was performed in a 33 nL volume.

The relative threshold cycle ( $C_{RT}$ ) value for each reaction was obtained using ExpressionSuite Software v1.0.4. Expression data ( $C_{RT}$  value) which had a standard deviation between the two technical replicates larger than 0.5 were removed. For the remainder, I used the mean of the  $C_{RT}$  values of the technical replicates for each individual. The expression level for each gene was normalized to *eflab* expression and the  $\Delta C_{RT}$  values ( $C_{RT\text{targeted gene}} - C_{RT\text{eflab}}$ ) were used for all downstream analyses. I used *eflab* as the endogenous reference gene because it has been shown to be an excellent endogenous control compared to other common reference genes in Atlantic salmon muscle tissue (Olsvik *et al.* 2005).



### Statistical analysis

To test for population effects on gene transcription between the two Atlantic salmon populations, I initially used the following model for each gene:

$$Y_{ijklmn} = \mu + P_i + T_j + S_k + D_l + I_m + e_{ijklmn} \quad (1)$$

where  $Y_{ijklmn}$  is the normalized transcription level (relative to the *eflab* gene;  $C_{RT\text{targeted gene}} - C_{RT\text{eflab}}$ ),  $P_i$  is the  $i^{\text{th}}$  effect of population;  $T_j$  is the  $j^{\text{th}}$  effect of tank;  $S_k$  is the  $k^{\text{th}}$  effect of sire;  $D_l$  is the  $l^{\text{th}}$  effect of dam;  $I_m$  is the  $m^{\text{th}}$  effect of interaction between sire and dam;  $e_{ijklmn}$  is the random residual. Population effect was considered as a fixed effect. Tank, sire, dam, and the interaction between sire and dam were random effects. I performed backward elimination to remove non-significant random effects from model (1) using the *step* function in the *lmerTest* R package (Kuznetsova, 2016). The final models (Supplementary Table S4.1) were used to determine whether there were significant population effects on gene transcription using the *anova* function in *lmerTest* (Kuznetsova *et al.* 2016).

To quantify the genetic architecture for all the genes within the two populations separately, I partitioned the total variance into sire, dam, and dam-by-sire interaction components using the following model:

$$Y_{ijklm} = \mu + T_i + S_j + D_k + I_l + e_{ijklm} \quad (2)$$

where tank (T), sire (S), dam (D), and the interaction between sire and dam (I) were treated as random effects. The significance for tank, sire, dam and the interaction were determined using the likelihood ratio test between the full model (model 2) and a reduced one without the tested effect using the *observLmer2* function in the *fullfact* R

package (Houde & Pitcher 2016). Gene transcription variance was partitioned into sire ( $V_S$ ), dam ( $V_D$ ), interaction between sire and dam ( $V_I$ ), and residual variance components in *fullfact*. The additive ( $V_A$ ), non-additive ( $V_{NA}$ ) and maternal variance ( $V_M$ ) were calculated as follows:  $V_A = 4V_S$ ;  $V_{NA} = 4V_I$ ;  $V_M = V_D - V_S$  (Lynch & Walsh 1998). To compare genetic architecture of gene transcription between the two populations across 22 genes, I conducted Wilcoxon signed-rank test for additive genetic, non-additive genetic, and maternal effects.

## RESULTS

### *Population effects on gene transcription*

I measured transcription at 22 genes using qRT-PCR in 426 Atlantic salmon fry. Nine genes showed significantly different transcription between the two populations (Figure 4.1; Supplementary Table S4.1). The *cpt2*, *myf5*, *myod1* and *tfam* genes showed significantly higher expression in the LaHave population while the *acadm*, *cs*, *ldha*, *mlc2* and *pygma* genes showed significantly higher expression in the Sebago population, with the expression difference ranging from 8% to 99% (Figure 4.1; Supplementary Table S4.1). While the difference in gene expression between the two populations for most analyzed genes is less than 30% expression difference, the expression of *ldha* and *pygma* in the Sebago population was close to double that of the LaHave population. Of the nine differentially expressed genes, five are involved in muscle energy generation; two genes are involved in muscle growth; one gene is involved in muscle energy regulation and one gene is involved in muscle contraction (Figure 4.1).

### *Significant effects on gene transcription within each population*

I tested for tank, sire, dam, and the interaction between sire and dam effects on gene transcription at all 22 genes within each population. Ten genes showed significant tank effects in the LaHave population and 11 genes showed significant tank effects in the Sebago population (Table 4.2). Seven genes showed significant tank effects in both populations. Among those seven genes, three genes are related to growth, two genes are related to muscle contraction and two genes encode enzymes involved in energy generation (Table 4.2). Within the LaHave population, five genes showed significant dam effects and two genes showed significant sire effects (Table 4.2). Within the Sebago population, four genes showed significant dam effects and one gene showed significant sire effects (Table 4.2). The *camk2g* gene showed a significant sire effect in both populations with similar levels of explained variance, while all the other significant sire or dam effects on gene expression were population-specific with different levels of explained variance between the two populations (Table 4.2). No genes showed significant sire-by-dam interaction effects in either population (Table 4.2). On average across all 22 genes, tank, sire, dam, and sire-by-dam interaction effects explained 13.0%, 4.2%, 7.3%, and 2.6% of the phenotypic variance in gene expression in the LaHave population, respectively, and 14.2%, 1.5%, 4.5%, and 4.2% of phenotypic variance in the Sebago population, respectively (Table 4.3).

### *Additive genetic, non-additive genetic, and maternal effects*

The two study populations exhibited both similarities and substantial differences in genetic variance components despite having been reared in a common environment

(Table 4.2; Figure 4.2). Maternal and genetic effects collectively explained 30.3% of the total phenotypic variance in the LaHave population and 25.8% of the total variance in the Sebago population. Specifically, in the LaHave population, additive genetic effects explained 16.7% of the total variance and non-additive genetic effects explained 10.5% of the total variance (Table 4.3). In the Sebago population, additive genetic effects explained 6.2% of the total phenotypic variance and non-additive genetic effects explained 16.6% of the total variance (Table 4.3). Maternal effects explained 3.1% of gene expression variance in the LaHave population; and 3.0% of gene expression variance in the Sebago population (Table 4.3).

I found 14 genes showed higher additive genetic effects in the LaHave population while four genes showed higher additive genetic effects in the Sebago population (Table 4.2; Figure 4.2). Seven genes showed higher non-additive genetic effects in the LaHave population and 10 genes showed higher non-additive genetic effects in the Sebago population (Table 4.2; Figure 4.2). Four genes showed no additive genetic effects and five genes showed no non-additive genetic effects in both populations. Half of the analyzed genes showed higher maternal effects in one population than the other population (Table 4.2; Figure 4.2). Across all 22 genes, the two populations were significantly different in additive genetic effects revealed by Wilcoxon signed-rank test ( $P = 0.01$ ), while there were no significant differences between the populations in non-additive genetic ( $P = 0.37$ ), and maternal effects ( $P = 0.73$ ).

## DISCUSSION

Population and individual differences in gene transcription have been reported in many studies (Oleksiak *et al.* 2002; Storey *et al.* 2007; Hutter *et al.* 2008). Although population-level variation was not the primary aim of this study, I found five of the seven metabolic enzyme genes involved in energy generation showed significant transcriptional differences between the two populations, after accounting for tank and family effects. Differences in the kinetic properties and concentrations of metabolic enzymes among populations and species are thought to be local adaptations to temperature, as they play an important role in homeostasis maintenance (Hochachka & Somero 1984; Crawford & Powers 1989; Crockett & Sidell 1990; Holland *et al.* 1997). Higher temperature tolerance is especially important for salmonids during migration to spawning sites (Eliason *et al.* 2011). However, my two study populations originate from habitats with similar average summer temperatures (Gradil 2015) and were reared under the same environmental conditions in the hatchery. The population-level differences in expression of those enzyme genes may be related to different thermal optimum as measured by Arrhenius breakpoint temperature (Gradil 2015) and this difference may underlie population differences in thermal tolerance.

Populations subject to different selection pressures can exhibit different distributions of genetic variance components, as, for example, strong selection reduces additive genetic variance (Carlson & Seamons 2008). In this study, I found the average additive genetic variance of transcription across the 22 genes associated with muscle function was higher in the LaHave population than in the Sebago population (Table 4.3). I expected the muscle function-related genes may have been under stronger selection in

the LaHave population than the Sebago population because the LaHave population originated from an anadromous (migratory) population while the Sebago population originated from a hatchery supplemented landlocked population with a shorter migration. Thus I predicted that the wild LaHave population would exhibit lower additive genetic variance for the transcription of those genes related to swimming performance and efficiency. However, the LaHave population fish have been in captivity for three generations while the Sebago population fish have been in the OMNRF hatchery for only one generation prior to this study. The longer captive rearing of the LaHave fish under relaxed selection pressures may have enabled the LaHave Atlantic salmon to recover additive genetic variance for the transcription of those genes.

The genetic architecture of phenotypic traits is important for organisms as it is the basis for their response to changing environments and underlies the nature and scope for evolutionary responses in new environments upon translocation. In this study, I partitioned the additive genetic, non-additive genetic, and maternal variance components of gene transcription using a full factorial breeding design. The average additive genetic variance component (16.7% in LaHave and 6.2% in Sebago) were comparable to studies in humans which reported a mean heritability ( $h^2$ ) of genome-wide gene transcription varied from 0.017 to 0.234 depending on the tissue (Price *et al.* 2011; Wright *et al.* 2014; Wheeler *et al.* 2016). My estimates were also comparable to a study of heritability of gene transcription in Atlantic salmon using a cDNA microarray consisting of 6 484 probes which found that heritability for most genes were low (Roberge *et al.* 2007). My estimates were lower than a study focused on transcription of four cytokine genes in Chinook salmon (*Oncorhynchus tshawytscha*) (mean  $h^2 = 0.256$ ) (Aykanat *et al.* 2012b)

and another study on transcription of three heat shock protein genes in sea turtles (*Caretta caretta*) (mean  $h^2 = 0.58$ ) (Tedeschi *et al.* 2016). The differences in additive genetic effects between my study and previous studies in other species may be due to different strengths of selection acting on the selected genes, the animals experiencing different evolutionary history or different experimental designs and analytical methods.

The observed differences in the magnitude of additive genetic effects for gene transcription across the functional groups is perhaps not surprising, as different traits can exhibit dramatic variation in heritability. For example, a review of  $h^2$  in salmonids found that 24 fitness-related traits exhibited a wide range of median heritabilities, varying from 0.02 to 0.51 (Carlson & Seamons 2008). In my study, I found genes involved in energy regulation generally had higher additive genetic components than genes involved in energy generation and muscle growth. This likely reflects different function genes having experienced different selection pressures; however, regardless of the mechanism, my results indicate that the functional groups will have different responses to selection in the short term. The majority of the genes included in my study exhibited lower heritabilities than the median heritability of fitness-related traits in salmonids (Carlson & Seamons 2008). This is surprising because fitness related traits are expected to show reduced additive genetic variance than gene expression as they are under strong selection (Mousseau & Roff 1987). It is possible that gene transcription may not follow classical quantitative genetic patterns. Nevertheless, the low additive genetic variance estimates suggest that the genes examined in the present study may be constrained in their evolutionary potential. However, a few genes in my study (e.g. *camk2g* and *tnnt1*) did exhibit high additive genetic variance, and may thus evolve more rapidly.

Although the additive genetic variance of gene transcription has been investigated in a variety of studies (e.g. Price *et al.* 2011; Aykanat *et al.* 2012b; Wright *et al.* 2014; Tedeschi *et al.* 2016; Wheeler *et al.* 2016), non-additive and maternal variance components for gene expression are seldom reported. In particular, dams usually have greater contributions to the phenotypes of their offspring at early life stages than sires because they not only contribute 50% of the nuclear genome, but they also influence offspring phenotype through cytoplasmic inheritance as well as classical maternal effects (Bernardo 1996; Wolf & Wade 2009). In my analyses, a greater number of genes showed significant dam effects than sire effects, and the dam variance component was higher than that of the sire. Aykanat *et al.* (2012b) also found that the dam variance component of gene transcription for four cytokine genes was higher than that of the sire in Chinook salmon. Videvall *et al.* (2016) found that the pattern of gene expression in hybrids between two *Arabidopsis lyrata* populations was more similar to the maternal population than to the paternal population. Generally, maternal effects are higher earlier in life and decrease with development (Heath *et al.* 1999), thus I expected to find higher maternal effects on transcription for the genes I surveyed at young juvenile life stage. However, in my study maternal effects were generally low and the average maternal effect was smaller than the average additive and non-additive genetic effects in both populations (Table 4.3), indicating maternal effects on gene transcription decreased more rapidly than fitness-related traits.



## CONCLUSIONS

I used Atlantic salmon from two populations with very different life histories in a full factorial breeding design to quantify gene transcription at 22 muscle function loci. I found that maternal and genetic effects combined explained 30.3% and 25.8% of the transcriptional variation across all assayed genes in the LaHave and Sebago populations, respectively. The two populations exhibited profound differences in genetic architecture of gene transcription at individual loci and among functional gene groups. Contrary to expectation for young life stages, the results also highlighted that maternal effects are lower than genetic effects for gene expression, at least for the genes examined. In addition, evolution by selection acting directly on gene expression is likely to be less effective because of its lower value of additive genetic effects than more traditional phenotypic traits.

## REFERENCES

- Aykanat T, Bryden CA, Heath DD (2012a) Sex-biased genetic component distribution among populations: additive genetic and maternal contributions to phenotypic differences among populations of Chinook salmon. *Journal of Evolutionary Biology*, **25**, 682–690.
- Aykanat T, Heath JW, Dixon B, Heath DD (2012b) Additive, non-additive and maternal effects of cytokine transcription in response to immunostimulation with *Vibrio* vaccine in Chinook salmon (*Oncorhynchus tshawytscha*). *Immunogenetics*, **64**, 691–703.
- Bernardo J (1996) Maternal effects in animal ecology. *American Zoologist*, **36**, 83–105.
- Buckland PR (2004) Allele-specific gene expression differences in humans. *Human Molecular Genetics*, **13**, R255–260.
- Carlson SM, Seamons TR (2008) A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. *Evolutionary Applications*, **1**, 222–238.
- Cheung VG, Spielman RS (2002) The genetics of variation in gene expression. *Nature Genetics*, **32 Suppl**, 522–525.
- Crawford DL, Powers DA (1989) Molecular basis of evolutionary adaptation at the lactate dehydrogenase-B locus in the fish *Fundulus heteroclitus*. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 9365–9369.
- Crawford DL, Powers DA (1992) Evolutionary adaptation to different thermal environments via transcriptional regulation. *Molecular Biology and Evolution*, **9**, 806–813.
- Crockett EL, Sidell BD (1990) Some pathways of energy metabolism are cold adapted in antarctic fishes. *Physiological Zoology*, **63**, 472–488.
- Desai BN, Myers BR, Schreiber SL (2002) FKBP12-rapamycin-associated protein associates with mitochondria and senses osmotic stress via mitochondrial dysfunction. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 4319–4324.
- Eliason EJ, Clark TD, Hague MJ *et al.* (2011) Differences in thermal tolerance among sockeye salmon populations. *Science*, **332**, 109–112.

- Evans ML, Neff BD (2009) Non-additive genetic effects contribute to larval spinal deformity in two populations of Chinook salmon (*Oncorhynchus tshawytscha*). *Aquaculture*, **296**, 169–173.
- Forsman A (2014) Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 302–307.
- Francetic T, Li Q (2011) Skeletal myogenesis and Myf5 activation. *Transcription*, **2**, 109–114.
- Fraser HB (2013) Gene expression drives local adaptation in humans. *Genome Research*, **23**, 1089–1096.
- Gallardo JA, Lhorente JP, Neira R (2010) The consequences of including non-additive effects on the genetic evaluation of harvest body weight in Coho salmon (*Oncorhynchus kisutch*). *Genetics Selection Evolution*, **42**, 19.
- Gleason LU, Burton RS (2015) RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail *Chlorostoma funebris*. *Molecular Ecology*, **24**, 610–627.
- Gradil KJH (2015) Thermal performance covaries with environmental temperature across populations of Atlantic salmon (*Salmo salar*). M.Sc. Thesis. The University of Western Ontario, Canada.
- Han R, Campbell KP (2007) Dysferlin and muscle membrane repair. *Current opinion in cell biology*, **19**, 409–416.
- Hansen MM, Olivieri I, Waller DM, Nielsen EE, GeM Working Group (2012) Monitoring adaptive genetic responses to environmental change. *Molecular Ecology*, **21**, 1311–1329.
- He X, Johansson ML, Heath DD (2016) Role of genomics and transcriptomics in selection of reintroduction source populations. *Conservation Biology*, **30**, 1010–1018
- He X, Wilson CC, Wellband KW *et al.* (2015) Transcriptional profiling of two Atlantic salmon strains: implications for reintroduction into Lake Ontario. *Conservation Genetics*, **16**, 277–287.
- Heath DD, Fox CW, Heath JW (1999) Maternal effects on offspring size: variation through early development of Chinook salmon. *Evolution*, **53**, 1605–1611.

- Hevrøy EM, Jordal A-EO, Hordvik I *et al.* (2006) Myosin heavy chain mRNA expression correlates higher with muscle protein accretion than growth in Atlantic salmon, *Salmo salar*. *Aquaculture*, **252**, 453–461.
- Hochachka PW, Somero GN (1984) *Biochemical Adaptation*. Princeton University Press, Princeton, NJ.
- Hodgins-Davis A, Townsend JP (2009) Evolving gene expression: from G to E to  $G \times E$ . *Trends in Ecology & Evolution*, **24**, 649–658.
- Holland LZ, McFall-Ngai M, Somero GN (1997) Evolution of lactate dehydrogenase-A homologs of barracuda fishes (genus *Sphyraena*) from different thermal environments: differences in kinetic properties and thermal stability are due to amino acid substitutions outside the active site. *Biochemistry*, **36**, 3207–3215.
- Houde ALS, Black CA, Wilson CC, Pitcher TE, Neff BD (2015) Genetic and maternal effects on juvenile survival and fitness-related traits in three populations of Atlantic salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, **72**, 751–758.
- Houde ALS, Pitcher TE (2016) fullfact: an R package for the analysis of genetic and maternal variance components from full factorial mating designs. *Ecology and Evolution*, **6**, 1656–1665.
- Houde AL, Wilson CC, Neff BD (2013) Genetic architecture of survival and fitness-related traits in two populations of Atlantic salmon. *Heredity*, **111**, 513–519.
- Hutter S, Saminadin-Peter SS, Stephan W, Parsch J (2008) Gene expression variation in African and European populations of *Drosophila melanogaster*. *Genome Biology*, **9**, R12.
- Jobling M (1993) Bioenergetics: feed intake and energy partitioning. In: *Fish Ecophysiology*. (eds Rankin JC, Jensen FB), pp. 1–44. Chapman and Hall, London.
- Jinks J, Perkins JM, Pooni H (1973) The incidence of epistasis in normal and extreme environments. *Heredity*, **31**, 263–269.
- King TL, Kalinowski ST, Schill WB, Spidle AP, Lubinski BA (2001) Population structure of Atlantic salmon (*Salmo salar* L.): a range-wide perspective from microsatellite DNA variation. *Molecular Ecology*, **10**, 807–821.
- King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science*, **188**, 107–116.
- Kuznetsova A, Brockhoff PB, Christensen RHB (2016) *lmerTest: tests in linear mixed effects models*.

- Lecker SH (2003) Ubiquitin-protein ligases in muscle wasting: multiple parallel pathways? *Current Opinion in Clinical Nutrition and Metabolic Care*, **6**, 271–275.
- López-Maury L, Marguerat S, Bähler J (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, **9**, 583–593.
- Løvoll M, Austbø L, Jørgensen JB, Rimstad E, Frost P (2011) Transcription of reference genes used for quantitative RT-PCR in Atlantic salmon is affected by viral infection. *Veterinary Research*, **42**, 8.
- Lynch M, Walsh B (1998) *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland, Massachusetts, USA.
- Mallona I, Weiss J, Egea-Cortines M (2011) pcrEfficiency: a Web tool for PCR amplification efficiency prediction. *BMC Bioinformatics*, **12**, 404.
- Mousseau TA, Roff DA (1987) Natural selection and the heritability of fitness components. *Heredity*, **59**, 181–197.
- Murakami M, Ichisaka T, Maeda M *et al.* (2004) mTOR is essential for growth and proliferation in early mouse embryos and embryonic stem cells. *Molecular and Cellular Biology*, **24**, 6710–6718.
- Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.
- Olsvik PA, Lie KK, Jordal A-EO, Nilsen TO, Hordvik I (2005) Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. *BMC Molecular Biology*, **6**, 21.
- Parrish DL, Behnke RJ, Gephard SR, McCormick SD, Reeves GH (1998) Why aren't there more Atlantic salmon (*Salmo salar*)? *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 281–287.
- Petretto E, Mangion J, Dickens NJ *et al.* (2006) Heritability and tissue specificity of expression quantitative trait loci. *PLoS Genetics*, **2**, e172.
- Price AL, Helgason A, Thorleifsson G *et al.* (2011) Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. *PLoS genetics*, **7**, e1001317.
- Roberge C, Guderley H, Bernatchez L (2007) Genomewide identification of genes under directional selection: gene transcription  $Q_{ST}$  scan in diverging Atlantic salmon subpopulations. *Genetics*, **177**, 1011–1022.

Schlichting CD, Smith H (2002) Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, **16**, 189–211.

Schoville SD, Barreto FS, Moy GW, Wolff A, Burton RS (2012) Investigating the molecular basis of local adaptation to thermal stress: population differences in gene expression across the transcriptome of the copepod *Tigriopus californicus*. *BMC Evolutionary Biology*, **12**, 170.

Schulte PM (2004) Changes in gene expression as biochemical adaptations to environmental change: a tribute to Peter Hochachka. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, **139**, 519–529.

Storey JD, Madeoy J, Strout JL *et al.* (2007) Gene-expression variation within and among human populations. *American Journal of Human Genetics*, **80**, 502–509.

Tedeschi JN, Kennington WJ, Tomkins JL *et al.* (2016) Heritable variation in heat shock gene expression: a potential mechanism for adaptation to thermal stress in embryos of sea turtles. *Proceedings of the Royal Society B*, **283**, DOI: 10.1098/rspb.2015.2320.

Torstensen B e., Nanton D a., Olsvik P a., Sundvold H, Stubhaug I (2009) Gene expression of fatty acid-binding proteins, fatty acid transport proteins (cd36 and FATP) and  $\beta$ -oxidation-related genes in Atlantic salmon (*Salmo salar* L.) fed fish oil or vegetable oil. *Aquaculture Nutrition*, **15**, 440–451.

Videvall E, Sletvold N, Hagenblad J, Ågren J, Hansson B (2016) Strong maternal effects on gene expression in *Arabidopsis lyrata* hybrids. *Molecular Biology and Evolution*, **33**, 984–994.

Wheeler HE, Shah KP, Brenner J *et al.* (2016) Survey of the heritability and sparsity of gene expression traits across human tissues. *bioRxiv*, 43653.

Whitehead A, Triant DA, Champlin D, Nacci D (2010) Comparative transcriptomics implicates mechanisms of evolved pollution tolerance in a killifish population. *Molecular Ecology*, **19**, 5186–5203.

Wolf JB, Wade MJ (2009) What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society B*, **364**, 1107–1115.

Wright FA, Sullivan PF, Brooks AI *et al.* (2014) Heritability and genomics of gene expression in peripheral blood. *Nature Genetics*, **46**, 430–437.

**Table 4.1** Quantitative real-time PCR primers and TaqMan probes for muscle function related genes (abbreviation in parentheses) used for Atlantic salmon (*Salmo salar*) genetic architecture analysis in two populations.

Gene name	NCBI number	TaqMan MGB probe	Primers
acyl-Coenzyme A dehydrogenase, long-chain ( <i>acadt</i> )	BT044691	ACAGGACACGGCTGAG	Fw: GCTGGAGAAGATTGGCCTGAT Rv: GACGCACGTCCTCAAAGAACA
acyl-Coenzyme A dehydrogenase, medium chain ( <i>acadm</i> )	NM_001139636	ACATTCCAGAGGACTGTGG	Fw: GAGCTGGGCCTGATGAAGTC Rv: CGAATATGCCCAGACCCATT
calcium/calmodulin-dependent protein kinase type II gamma chain ( <i>camk2g</i> )	EG869390	CAGTGGTGCGCAGAT	Fw: GCTCGGGAAGGGAGCTTTT Rv: TCCTGACCTGATGACTTCTTCACA
carnitin palmitoyltransferase 2 ( <i>cpt2</i> )	BG934647	TGGGCTACGGTGTCC	Fw: GGTGCCCCGATGGATTTCG Rv: TGCAGCCAATCCACTCATCA
citrate synthase ( <i>cs</i> )	DY741160	CTGGCTAACCAGGAGGT	Fw: TGGCTGGACCCCTTCATG Rv: GGGCCGTCAACCATACCA
dihydrolipoamide S-acetyltransferase ( <i>dlat</i> )	DY740452	TTTGACGTGGCCAGCAT	Fw: GCCTGCTGACAATGAGAAAGG Rv: CACAACCAACGTCACCGACAT
dysferlin-interacting protein 1 ( <i>dysi1</i> )	NM_001146538	CTGACATTGCCAAGTAC	Fw: GGCCTGCAGTGACGGATT Rv: GGTGCGCACCAATAGAAAGC
elongation factor 1A ( <i>ef1ab</i> )*	BG933853	AAATCGGCGGTATTGG	Fw: TGCCCTCCAGGATGTCTAC Rv: CACGGCCCCACAGGTACTG
muscle fatty acid binding protein mRNA ( <i>fabp3</i> )*	AY509548	TCAAGTCCCTAATAACG	Fw: CACCGCTGACGACAGGAAA Rv: TGCACGTGAACCATCTTACCA
FKBP12-rapamycin complex-associated protein ( <i>frap1</i> )	EG909867	CTAGCAAATAACCAGGGCC	Fw: GCCAGTGCCTTGAGCAATAAG Rv: CGATGGCTTTGGGAAACG
fast myosin heavy chain ( <i>fmyhc</i> )*	BE518566	CCACTGAAAACAAGGTAAAA	Fw: CCAAAGTGGAGAAGGAGAAGCA Rv: CATTGACGCCATCTCCTCTGT
lactate dehydrogenase A4 ( <i>ldha</i> )	NM_001139642	TGGTCTGACCGACGTC	Fw: ATGCGTGCTGGGCAACA Rv: CGGGCTTCAGGGTCATGT
myosin regulatory light chain 2 ( <i>mlc2</i> )	NM_001123716	CCGTCTTCCTCACCATG	Fw: GCGGCCCCATCAACTTC Rv: CACCCTTGAGCTTCTCTCCAA
muscle RING finger 1 ( <i>murf1</i> )	NM_001279122	CCTGCCCTGCCAAC	Fw: TGTTCCAGAAGCCCGTAGTCA Rv: AGCCGCGGCACAGGTT
myogenic regulatory factor 5 ( <i>myf5</i> )	DQ452070	CGCAACGCCATCCA	Fw: GCTGCCTAAGGTGGAGATCCT Rv: CTCCTGGAGGCTCTCGATGT
myoblast determination protein 1 ( <i>myod1</i> )	NM_001123601	TGCAAGAGGAAAACC	Fw: CCTCTGGGCATGCAAAGC Rv: TTCCTCCGGTCGGTGTG

peroxisome proliferator-activated receptor alpha ( <i>ppara</i> )	NM_001123560	ACGGTCACAGAGCTA	Fw: GCCAGTGCACCTCCGTAGA Rv: GGGACAGACTTGGCGAACTC
5'-AMP-activated protein kinase subunit gamma-2 ( <i>prkag2</i> )	EG800235	CTCAGCCTTAATTATG	Fw: CACTAGGACCAGCCGATGGA Rv: GACCCTGGAGACCATCATTGA
phosphorylase, glycogen (muscle) A ( <i>pygma</i> )	NM_001139650	AAGCTGTGCGTCCAC	Fw: CCCCAGATGAGCAACTCAAA Rv: AACTGGTGAAGAAGGGAAGTATATGG
slow myosin heavy chain 1 ( <i>smyh1</i> )	DQ369355	CTGCTGTGCTGTTTAA	Fw: GATGTTTACCTTCCTGCATGAG Rv: GGCTGCGTAACGCTCTTTG
mitochondrial transcription factor A ( <i>tfam</i> )	BT048987	CCCGGTCGTCCTTTA	Fw: TCTGGGCAAACCCAAACGT Rv: CAAAGTGTTCTGCCATGAAGATG
troponin I, fast skeletal muscle ( <i>ttni2</i> )	BT048139	ACTGGCGTAAGAACAT	Fw: GGAATTGCGTGACGTTGGT Rv: CGTCCATACCGGCCTTGTC
troponin T, slow skeletal muscle ( <i>tnt1</i> )	BT057444	AGTCAGCGATCATCA	Fw: ATGATGTCACCGTACTCAGGAATC Rv: TCCTGGTCCCCTTGGTAACTT

\* The sequences of primers and probes for three genes were obtained from journal publications: *eflab* (Løvøll *et al.*, 2011), *fabp3* (Torstensen *et al.*, 2009) and *fmyhc* (Hevrøy *et al.*, 2006). The other primers and probes were designed for this study.



**Table 4.2** Summary of the genetic variance components of gene transcription for 22 genes associated with muscle function (arranged by putative function) in juvenile Atlantic salmon (*Salmo salar*) from two populations (LaHave and Sebago).

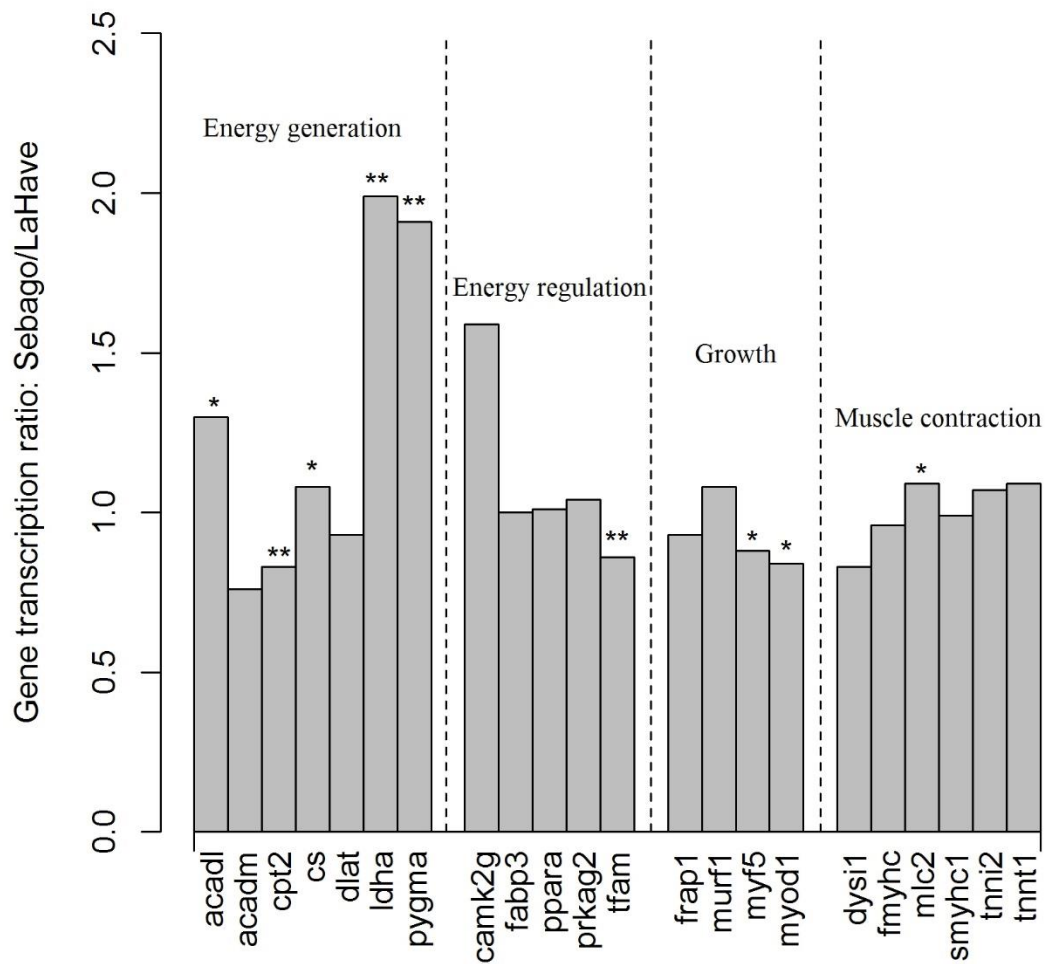
Gene	Population	V <sub>T</sub> (%)	V <sub>I</sub> (%)	V <sub>S</sub> (%)	V <sub>D</sub> (%)	V <sub>A</sub> (%)	V <sub>NA</sub> (%)	V <sub>M</sub> (%)
<b>Energy generation</b>								
<i>acadl</i>	LaHave	7.9	0.0	0.0	4.0	0.0	0.0	4.0
	Sebago	<b>15.1</b>	0.0	0.0	<b>10.6</b>	0.0	0.0	10.6
<i>acadm</i>	LaHave	3.1	3.2	7.7	<b>33.7</b>	30.7	12.9	26.1
	Sebago	4.6	5.5	0.0	3.7	0.0	22.1	3.7
<i>cpt2</i>	LaHave	<b>16.0</b>	0.0	0.0	2.1	0.0	0.0	2.1
	Sebago	<b>24.7</b>	8.5	0.0	3.7	0.0	34.2	3.7
<i>cs</i>	LaHave	10.7	0.6	0.0	8.0	0.0	2.5	8.0
	Sebago	10.8	0.0	0.0	0.0	0.0	0.0	0.0
<i>dlat</i>	LaHave	<b>25.3</b>	4.4	3.0	1.6	11.8	17.6	-1.4
	Sebago	<b>43.8</b>	9.9	0.0	1.7	0.0	39.4	1.7
<i>ldha</i>	LaHave	11.5	0.5	0.8	<b>17.3</b>	3.2	1.8	16.5
	Sebago	<b>16.0</b>	4.2	0.5	3.9	2.1	16.6	3.4
<i>pygma</i>	LaHave	10.0	3.4	3.1	<b>19.7</b>	12.3	13.5	16.6
	Sebago	<b>15.1</b>	4.5	0.0	8.2	0.0	18.1	8.2
<b>Energy regulation</b>								
<i>camk2g</i>	LaHave	0.0	8.8	<b>19.7</b>	6.0	78.9	35.0	-13.7
	Sebago	3.5	0.0	<b>13.7</b>	<b>9.1</b>	54.9	0.0	-4.7
<i>fabp3</i>	LaHave	<b>15.8</b>	0.0	0.0	<b>10.7</b>	0.0	0.0	10.7
	Sebago	3.1	0.0	1.6	0.0	6.3	0.0	-1.6
<i>ppara</i>	LaHave	11.1	0.0	0.0	1.4	0.0	0.0	1.4
	Sebago	<b>20.8</b>	5.6	0.0	0.0	0.0	22.3	0.0
<i>prkag2</i>	LaHave	<b>17.3</b>	5.0	1.5	4.7	6.0	19.8	3.2
	Sebago	3.3	15.4	0.0	0.0	0.0	61.6	0.0
<i>tfam</i>	LaHave	10.8	0.0	0.0	2.6	0.0	0.0	2.6
	Sebago	9.5	17.2	4.7	2.9	18.9	68.6	-1.8
<b>Growth</b>								
<i>frap1</i>	LaHave	4.3	8.9	2.4	5.6	9.5	35.8	3.3
	Sebago	4.0	0.0	0.0	0.0	0.0	0.0	0.0
<i>murf1</i>	LaHave	<b>29.5</b>	0.0	0.6	<b>12.1</b>	2.6	0.0	11.5
	Sebago	<b>15.2</b>	0.0	1.4	2.3	5.6	0.0	0.9
<i>myf5</i>	LaHave	<b>23.9</b>	0.3	5.1	9.8	20.4	1.1	4.7
	Sebago	<b>35.5</b>	0.0	0.0	0.0	0.0	0.0	0.0
<i>myod1</i>	LaHave	<b>15.6</b>	0.0	0.4	3.3	1.7	0.0	2.9
	Sebago	<b>16.4</b>	8.4	0.0	6.1	0.0	33.7	6.1
<b>Muscle contraction</b>								
<i>dysi1</i>	LaHave	<b>19.6</b>	3.3	3.5	0.0	14.1	13.3	-3.5
	Sebago	7.3	0.0	5.2	<b>21.9</b>	20.8	0.0	16.7
<i>fmyhc</i>	LaHave	8.7	9.5	3.1	2.3	12.4	38.1	-0.8
	Sebago	10.0	3.7	0.0	5.7	0.0	14.9	5.7
<i>mlc2</i>	LaHave	0.7	10.1	2.7	4.1	10.7	40.5	1.4
	Sebago	4.3	1.9	0.0	3.1	0.0	7.7	3.1
<i>smyhc1</i>	LaHave	<b>17.3</b>	0.0	6.9	3.4	27.5	0.0	-3.5
	Sebago	<b>22.2</b>	0.0	0.0	3.3	0.0	0.0	3.3
<i>tnni2</i>	LaHave	<b>20.4</b>	0.0	7.0	6.7	27.8	0.0	-0.3
	Sebago	<b>16.7</b>	6.5	1.0	3.6	3.9	26.2	2.6
<i>tnnt1</i>	LaHave	6.5	0.0	<b>24.5</b>	1.3	98.1	0.0	-23.2
	Sebago	10.9	0.0	5.7	<b>9.1</b>	22.9	0.0	3.4

Note: Significant tank, sire and dam effects ( $P < 0.05$ ) are marked with boldface type. The sire-dam interaction effects were not significant.  $V_T$ ,  $V_I$ ,  $V_S$ , and  $V_D$  represent the percentage of gene expression variance explained by tank, sire-dam interaction, sire, and dam effects, respectively.  $V_A$ ,  $V_{NA}$ , and  $V_M$  represent additive genetic, non-additive genetic, and maternal effects, respectively.

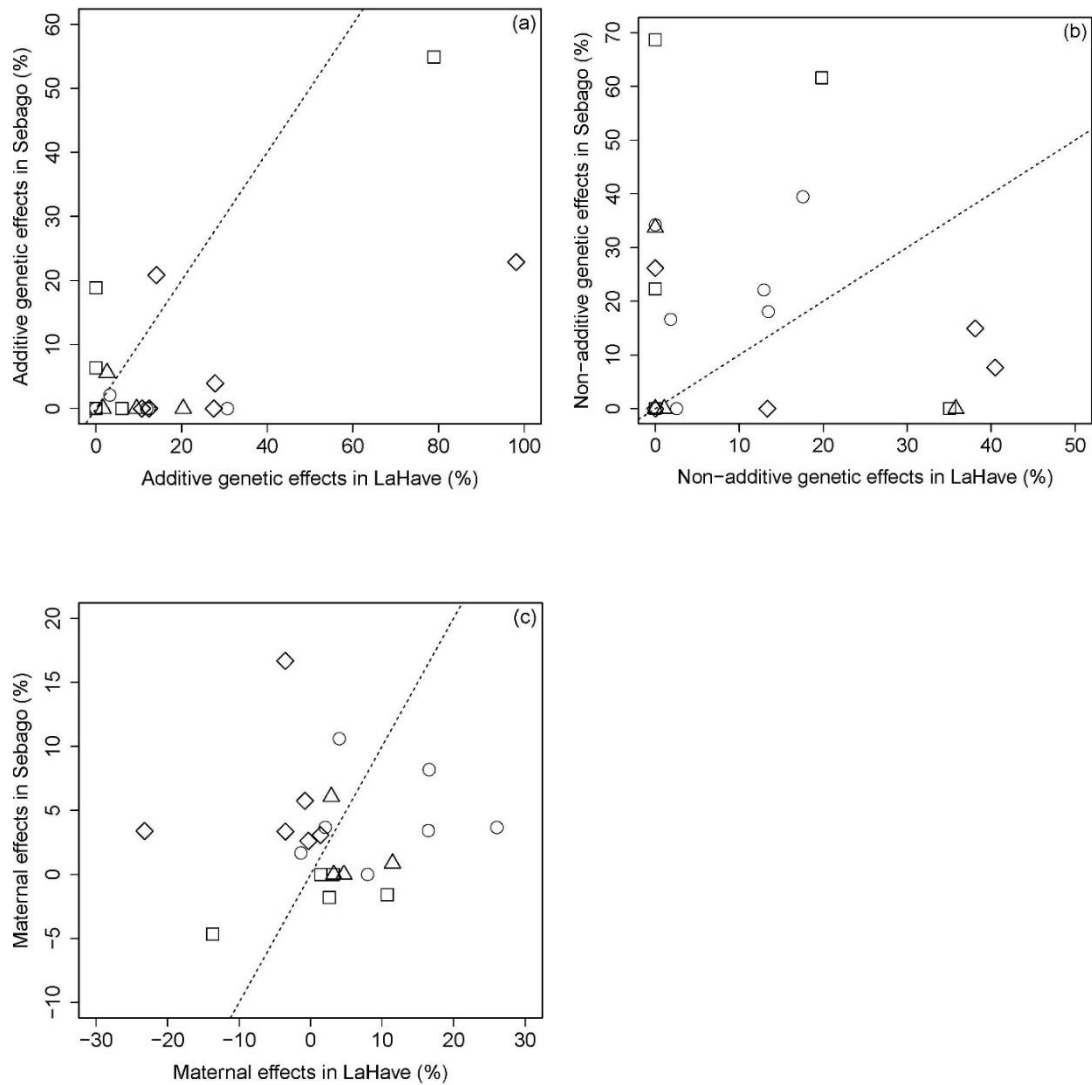
**Table 4.3** Summary of genetic architecture for gene transcription of four functional categories across 22 muscle-function related genes for juvenile Atlantic salmon (*Salmo salar*) from two populations (LaHave and Sebago).

Gene category	Population	V <sub>T</sub> (%)	V <sub>I</sub> (%)	V <sub>S</sub> (%)	V <sub>D</sub> (%)	V <sub>A</sub> (%)	V <sub>NA</sub> (%)	V <sub>M</sub> (%)
Energy generation	LaHave	12.1	1.7	2.1	12.3	8.3	6.9	10.3
	Sebago	18.6	4.7	0.1	4.5	0.3	18.6	4.5
Energy regulation	LaHave	11.0	2.7	4.2	5.1	17.0	11.0	0.8
	Sebago	8.0	7.6	4.0	2.4	16.0	30.5	-1.6
Growth	LaHave	18.3	2.3	2.1	7.7	8.5	9.2	5.6
	Sebago	17.8	2.1	0.3	2.1	1.4	8.4	1.7
Muscle contraction	LaHave	12.2	3.8	7.9	3.0	31.8	15.3	-5.0
	Sebago	11.9	2.0	2.0	7.8	7.9	8.1	5.8
All genes	LaHave	13.0	2.6	4.2	7.3	16.7	10.5	3.1
	Sebago	14.2	4.2	1.5	4.5	6.2	16.6	3.0

Note: V<sub>T</sub>, V<sub>I</sub>, V<sub>S</sub>, and V<sub>D</sub> represent the percentage of gene expression variance explained by tank, sire-dam interaction, sire, and dam effects, respectively. V<sub>A</sub>, V<sub>NA</sub>, and V<sub>M</sub> represent additive genetic, non-additive genetic, and maternal effects, respectively.



**Figure 4.1** Differences in gene transcription between two Atlantic salmon (*Salmo salar*) populations (LaHave and Sebago) measured using qRT-PCR. Gene names are described in Table 4.1. \*  $P < 0.05$ ; \*\* $P < 0.01$ .



**Figure 4.2** Scatterplots showing the distribution of muscle function gene transcription variance components for two Atlantic salmon (*Salmo salar*) populations (LaHave and Sebago). Comparisons of variance components between populations: (a) additive genetic effects, (b) non-additive genetic effects, and (c) maternal effects. Each symbol represents one gene. Symbols are coded for the four different functional gene groups: circles (○) represent energy generation genes; squares (□) represent genes involved in energy regulation; triangles (△) represent growth genes; diamonds (◇) represent muscle contraction genes. The dashed line represents the expected 1:1 relationship (i.e.,  $y = x$ ) between the variance components of the two study populations. Points above the 1:1 line indicate higher values in the Sebago population and points below the dashed line indicate higher values in the LaHave population.

**Supplementary Table S4.1** Model selection and population effects for the 22 analyzed genes in the two populations of Atlantic salmon (*Salmo salar*).

Gene	Final model	Population effect <i>P</i> value	Expression ratio (Sebago/LaHave)
<b>Energy generation</b>			
<i>acadl</i>	Population + Tank + Dam	<b>0.014</b>	1.30
<i>acadm</i>	Population + Dam + Dam:Sire	0.081	0.76
<i>cpt2</i>	Population + Tank	<b>0.002</b>	0.83
<i>cs</i>	Population + Tank	<b>0.015</b>	1.08
<i>dlat</i>	Population + Tank	0.200	0.93
<i>ldha</i>	Population + Tank + Dam	<b>&lt;0.001</b>	1.99
<i>pygma</i>	Population + Tank + Dam	<b>0.001</b>	1.91
<b>Energy regulation</b>			
<i>camk2g</i>	Population + Dam + Sire	0.082	1.59
<i>fabp3</i>	Population + Tank + Dam	0.949	1.00
<i>ppara</i>	Population + Tank	0.894	1.01
<i>prkag2</i>	Population + Tank + Dam:Sire	0.536	1.04
<i>tfam</i>	Population + Tank + Dam:Sire	<b>&lt;0.001</b>	0.86
<b>Growth</b>			
<i>frap1</i>	Population + Tank	0.072	0.93
<i>murfl</i>	Population + Tank + Dam	0.435	1.08
<i>myf5</i>	Population + Tank	<b>0.022</b>	0.88
<i>myod1</i>	Population + Tank + Dam	<b>0.012</b>	0.84
<b>Muscle contraction</b>			
<i>dysi1</i>	Population + Tank + Dam + Sire	0.383	0.83
<i>fmyhc</i>	Population + Tank+ Dam:Sire	0.324	0.96
<i>mlc2</i>	Population + Dam:Sire	<b>0.030</b>	1.09
<i>smyhc1</i>	Population + Tank	0.772	0.99
<i>tnni2</i>	Population + Tank + Dam + Sire	0.467	1.07
<i>tnnt1</i>	Population + Tank + Dam+ Sire	0.497	1.09

## CHAPTER 5

# TRANSCRIPTOME RESPONSE OF ATLANTIC SALMON TO COMPETITION WITH ECOLOGICALLY SIMILAR NON-NATIVE SPECIES<sup>4</sup>

## INTRODUCTION

The wide establishment of non-native species is one of the major global environmental challenges caused by human activities. It is estimated that there are close to a half million exotic species introduced in different ecosystems (Pimentel *et al.* 2001). Introduced species can provide conservation values to local ecosystems. For example, non-native plants can provide habitat for native species (Severns & Warren 2008; Sogge *et al.* 2008); non-native animals (e.g. crayfish and round goby) can be food sources for threatened native species and increase the number of threatened species (King *et al.* 2006; Tablado *et al.* 2010). However, introduced species more often threaten local biodiversity through pathogen introduction, predation and competition (Manchester & Bullock 2000; McDowall 2003; Vitule *et al.* 2009; Peeler *et al.* 2010). Among the negative effects that non-native species bring, competition with non-native species for limited resources in the ecosystem may attract the least attention because it is not a common cause of extirpation of native species (Davis 2003) and its negative impact on native species is not consistent (Turek *et al.* 2013).

Indeed, the presence of introduced species can affect growth, reproduction and survival of ecologically similar native species, presumably due to interspecific

---

<sup>4</sup> This is the outcome of joint research.

competition (Scott *et al.* 2003; Houde *et al.* 2015a, 2016). Interspecific competition affects low status species disproportionately (Gilmour *et al.* 2005), and competition stress, like any form of stress will have negative effects on the stressed species, including decreased growth, loss of immune function and reduced survival (Barton 2002; Gilmour *et al.* 2005). Because introduced non-native species act as biotic stressors, they may not only drive declines in native species, but may also be a significant barrier to the reintroduction of locally extirpated native species.

The presence of established non-native salmonids in Lake Ontario, Canada, has been proposed as one of the reasons for the unsuccessful reintroduction of Atlantic salmon (*Salmo salar*) into Lake Ontario (Jones & Stanfield 1993; Scott *et al.* 2005). Atlantic salmon was a native species in Lake Ontario until extirpated in the late 1800's, and decades of reintroduction efforts have been largely unsuccessful (Dimond & Smitka 2005). A number of non-native salmonid species (Chinook salmon *Oncorhynchus tshawytscha*, coho salmon *O. kisutch*, rainbow trout *O. mykiss* and brown trout *S. trutta*) have been introduced into Lake Ontario to address the recreational need for large salmonid sport fish (Stewart & Schaner 2002). Some of the established non-native species, such as rainbow trout and brown trout, are more aggressive than Atlantic salmon (Van Zwol *et al.* 2012) and thus represent potentially highly stressful competitors. While it is generally difficult to demonstrate effects of interspecific competition in the wild (Hastings 1987), experiments in artificial streams are a good alternative because the competitively limited resources can be largely controlled, and the effects caused by interspecific competition can be quantified. Studies in artificial stream tanks revealed that competition with juvenile rainbow and brown trout has negative effects on the growth



and survival of juvenile Atlantic salmon (Houde *et al.* 2015a; b). It was also reported that the presence of rainbow trout in natural streams reduces the growth of Atlantic salmon (Houde *et al.* 2016). Furthermore, although juvenile Chinook salmon were found to have no negative effects on juvenile Atlantic salmon growth or survival (Houde *et al.* 2015a), adult Chinook salmon can affect the survival of adult Atlantic salmon during reproduction (Scott *et al.* 2003). Those studies demonstrated that interspecific competition between Atlantic salmon and non-native salmonids can affect the establishment of Atlantic salmon after release in Lake Ontario, but the mechanisms that mediate the negative effects on Atlantic salmon at the molecular level are largely unknown.

Gene expression is the process whereby genetic information stored in genome is used to synthesize functional products and hence determine phenotype. Gene expression is determined by both genetic and environmental factors (Buckland 2004; Petretto *et al.* 2006; López-Maury *et al.* 2008; Hodgins-Davis & Townsend 2009). Changes in gene expression are the mechanisms behind acclimation and adaption to environmental stress (Schulte 2004). Gene expression changes in response to abiotic stress have been widely studied in fish species, and those studies deepened our understanding of population differences in response to, and tolerance of, thermal stress (e.g. Narum & Campbell 2015), pollution exposure (e.g. Whitehead *et al.* 2010), and salinity (e.g. Brennan *et al.* 2015). While most studies on biotic stress in fish focus on immune challenge (e.g. Wellband & Heath 2013) and individuals with different social rank (e.g. Trainor & Hofmann 2007; Schunter *et al.* 2014), studies of the transcriptional response to interspecific competition due to niche overlap have not been reported.

In this study, I conducted controlled interspecific competition experiments between Atlantic salmon and three ecologically similar non-native salmonids (Chinook salmon, rainbow trout, and brown trout) which are established in Lake Ontario. I examined transcriptional response to competition in two Atlantic salmon populations (LaHave and Sebago) with very different life histories (anadromous and landlocked) which likely have different competitive ability. Houde *et al.* (2015a; b) conducted interspecific competition trials between these same Atlantic salmon populations and four non-native salmonids in artificial stream tanks to investigate effects of interspecific competition on growth and survival of Atlantic salmon. The results showed that the Sebago population exhibited overall faster growth and higher survival than the LaHave population when reared with rainbow trout and brown trout in two replicate studies (Houde *et al.* 2015a; b). I included four treatments for each Atlantic salmon population: Atlantic salmon reared alone and Atlantic salmon reared with each of the three non-native salmonids. After 10 months of rearing in artificial stream tanks, I used RNA-Seq to compare the transcriptome of the two Atlantic salmon populations in response to interspecific competition. I found population-specific response to competition with non-native species, highlighting issues with the selection of the source for reintroduction efforts. My results also implied the difficulty in predicting the effect of introduced species on native species, as local evolutionary history can result in very different response patterns.

## MATERIALS AND METHODS

### *Design and sampling*

Two Atlantic salmon populations were used in this study: LaHave and Sebago. For each Atlantic salmon population, I created four treatments: Atlantic salmon reared alone and Atlantic salmon reared with one of three salmonids (Chinook salmon, rainbow trout, and brown trout). All fish used in this study were provided by the Ontario Ministry of Natural Resources (OMNR), Canada. Atlantic salmon eggs from both populations were fertilized in November 2011 and transferred to artificial stream tanks in September 2012. Detailed information of the design of the artificial stream tanks are provided in Houde *et al.* (2015a). Initially, there were 32 Atlantic salmon in each of the tanks where Atlantic salmon were reared alone, and there were 16 Atlantic salmon and 16 fish of the competing species in each of the tanks where Atlantic salmon were reared with the non-native species. After 10 months in the artificial stream tanks, Atlantic salmon were humanely euthanized using an overdose of buffered MS-222. I collected spleens from the juvenile Atlantic salmon and stored them in RNAlater. I chose to sample spleen tissue for this study because the spleen is sensitive to whole-organism stress and is associated with circulating blood cells and immune response (Peters & Schwarzer 1985; Hernandez *et al.* 2013).

### *RNA isolation*

RNA was extracted from spleen tissue using Trizol (Invitrogen, California, United States) following the manufacture's protocol. The quality and concentration of RNA was checked using Agilent RNA 6000 Nano Kit in an Agilent 2100 Bioanalyzer (Agilent Technologies, Mississauga, ON, Canada). I selected RNA samples with RNA integrity

number (RIN) greater than 7.0 from four fish taken from the same competition treatment and combined them using equal amounts of total RNA. The mixed RNA samples were treated for possible genomic DNA contamination using TURBO™ DNase (Invitrogen, California, United States). After DNase treatment, the quality and concentration of the RNA samples were checked again using the Agilent 2100 Bioanalyzer (Agilent Technologies, Mississauga, ON, Canada) and sent to BGI Americas Corporation for RNA sequencing. For Atlantic salmon reared alone, I sent two separate pooled RNA samples for each population. For Atlantic salmon reared with each of the other three species, I sent one pooled RNA sample for each population. In total, 10 pooled RNA samples were sent and RNA sequencing was performed on the Illumina HiSeq™ 2000 platform.

### *Data analysis*

To obtain sequence read counts for each gene, I followed the protocol described in Anders *et al.* (2013). Briefly, clean reads were mapped to the Atlantic salmon genome (NCBI accession no.: AGKD000000000.3) using Bowtie 1.1.1 (Langmead *et al.* 2009) and Tophat 2.0.13 (Trapnell *et al.* 2012). Then, I used samtools 1.2 (Li *et al.* 2009) to sort and create the SAM files. After that, I used HTSeq-0.6.1 (Anders *et al.* 2015) to count reads for each gene. The principle component analyses and the construction of the distance heatmap were performed using DESeq2 (Love *et al.* 2014). I used Cufflinks to obtain the mRNA sequences of each gene (Trapnell *et al.* 2012). The longest isoform of each gene was extracted and used for blastx search in the non-redundant protein database using

blast 2.2.30 (Camacho *et al.* 2009). The results obtained from local blastx were loaded into Blast2GO (Conesa *et al.* 2005) for GO term mapping and annotation.

To test for transcriptomic differences between the two Atlantic salmon populations in response to competition with the three different non-native salmonid species within each Atlantic salmon population, I used GFOLD V1.1.3 because GFOLD can analyze RNA-Seq data with biological replicates or one treatment without replicates with a reliable statistical approach (Feng *et al.* 2012). To prepare gene expression data for GFOLD, I followed GFOLD's manual to obtain reads per kilobase of transcript per million mapped reads (RPKM) value for each gene in each sample. To identify differentially expressed genes by GFOLD, c value was set to 0.01 as default and genes with GFOLD value larger than 1 or less than -1 were accepted as significantly differently expressed. To evaluate population differences, I analyzed differentially expressed genes between the LaHave the Sebago Atlantic salmon reared alone samples. To quantify transcriptomic response to interspecific competition, I compared the Atlantic salmon reared with one non-native species to the two control samples (Atlantic salmon alone) within each population. The functional categorization of significantly differentially expressed genes in response to interspecific competition was plotted using BGI WEGO (Ye *et al.* 2006).

#### *Quantitative real-time PCR*

To validate RNA-Seq results, I designed primers for 14 genes (Supplementary Table S5.1) which showed significant differences in at least one of the six comparisons among four samples: LaAS1 (LaHave Atlantic salmon reared alone sample 1), LaBT

(LaHave Atlantic salmon reared with brown trout), SeAS1 (Sebago Atlantic salmon reared alone sample 1) and SeBT (Sebago Atlantic salmon reared with brown trout) as identified by GFOLD. I used *ribosomal protein S20 (rps20)* as the endogenous control as it has been shown to be invariant in Atlantic salmon spleen (Olsvik *et al.* 2005). I measured gene expression in three individuals from each of four treatments: LaHave reared alone, LaHave reared with brown trout, Sebago reared alone, and Sebago reared with brown trout. For each fish, I had three technical replicates. TURBO™ DNase treated RNA was used for cDNA synthesis using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Burlington, ON, Canada). The cDNA samples were diluted 1:10 for qRT-PCR analysis. The qRT-PCR reactions were conducted in 10 µL reactions which consisted of 5 µL SYBR Select Master Mix (Applied Biosystems, Burlington, ON, Canada), 0.5 µL 10 mM primers and 1 µL diluted cDNA. The qRT-PCR was performed in a QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems, Burlington, ON, Canada). The relative expression of each targeted gene was normalized to the expression of *rps20*.

## RESULTS

### *Sequencing summary and reads mapping*

The RNA-Seq data have been submitted to the NCBI SRA database (SRA accession: SRP080309). In total, I obtained approximately 160 million 90 bp paired-end high quality (Q>20) clean reads from the 10 pooled RNA samples. The number of sequence per sample ranged from 14.9 million to 16.9 million. Overall, 78.8% of the

reads mapped to the Atlantic salmon draft genome and the mapping rate for each sample varied from 72.9% to 81.3% (Supplementary Table S5.2).

#### *Principle component analysis and distance heatmap*

Principle component analysis based on gene transcription of the 2 000 genes which had the highest variation among samples showed that PC1 and PC2 explained 42% and 25% of the variance respectively (Figure 5.1a). The two populations were separated by PC1, while PC2 primarily reflected variation among competition treatments (Figure 5.1a). However, clear population-specific responses to competition were evident. For example, competition with rainbow trout resulted in the largest transcriptional response in the LaHave population while competition with brown trout resulted in the largest transcriptional response in the Sebago population (Figure 5.1a).

Within each population, Atlantic salmon responded transcriptionally different to competition with the three introduced salmonid species (Figures. 5.1b and 5.1c). The LaHave Atlantic salmon reared with Chinook salmon showed a different transcriptome response compared to the LaHave Atlantic salmon reared with rainbow trout along both PC1 and PC2 (Figure 5.1b). The Sebago Atlantic salmon reared with Chinook salmon showed a similar transcriptome response to the Sebago Atlantic salmon reared with rainbow trout, and the Sebago Atlantic salmon under these two treatments showed different transcriptome response compared to the Sebago reared with brown trout (Figure 5.1c).

In the distance heatmap, the five samples within each population clustered together, reflecting the large population effect on the transcriptome (Figure 5.2). Within

the LaHave population, the Atlantic salmon reared alone and Chinook salmon competition samples clustered together, while the LaHave Atlantic salmon reared with rainbow trout and brown trout clustered together (Figure 5.2). Within the Sebago population, the two Atlantic salmon reared alone samples clustered together while the Atlantic salmon reared with Chinook salmon and with rainbow trout clustered together (Figure 5.2). The Atlantic salmon reared with brown trout showed a highly divergent transcriptional profile within the Sebago population (Figure 5.2). The distance heatmap and PCA analyses indicated that the effects of population on gene expression were higher than that of interspecies competition.

#### *Gene expression differences*

GFOLD showed that 266 genes were transcribed at significantly different levels between the two Atlantic salmon populations when reared alone. Within the LaHave population, there were 209, 350, and 701 genes that exhibited a significant response to competition with Chinook salmon, brown trout, and rainbow trout, respectively (Figure 5.3). Within the Sebago population, there were 131, 384, and 191 genes that responded to competition with Chinook salmon, brown trout, and rainbow trout, respectively (Figure 5.3). Within LaHave, there were 10 genes that exhibited a significant response among all three interspecific competition treatments (Supplementary Figure S5.1a; Supplementary Table S5.3). Within Sebago, there were 9 genes that responded among all three interspecific competition treatments (Supplementary Figure S5.1b; Supplementary Table S5.4).



There were only 23, 13, and 19 genes shared between the two populations in response to competition with Chinook salmon, brown trout, and rainbow trout, respectively. Among the 23 genes showing a common response to competition with Chinook salmon in both populations, 20 showed the same trend of regulation of gene expression (Supplementary Table S5.5). Among the 13 genes showing a common response to competition with brown trout in both populations, 3 showed the same trend of regulation (Supplementary Table S5.6). Among the 19 genes showing a common response to competition with rainbow trout in both populations, 6 showed the same trend of regulation (Supplementary Table S5.7). While most responding genes were population-specific, the GO term analysis using the combined responding genes within each population showed that the responding genes were involved in similar functional groups in both populations (Supplementary Figure S5.2).

#### *Comparison between gene expression level revealed by qRT-PCR and RNA-seq*

I quantified 14 genes in 12 fish from four treatments (LaHave reared alone, LaHave reared with brown trout, Sebago reared alone, and Sebago reared with brown trout). The spearman correlation coefficient between relative expression quantified by qRT-PCR and RNA-Seq was 0.81 (Supplementary Figure S5.3).

## DISCUSSION

The establishment of non-native species can negatively affect the fitness of less aggressive native species (Fausch 2007; Turek *et al.* 2013). While gene expression response to many environmental stresses have been investigated (Whitehead *et al.* 2010;

Wellband & Heath 2013; Narum & Campbell 2015; Brennan *et al.* 2015), to my knowledge, transcriptional responses to competition with ecologically similar species has not been reported in fish. In this study, I used RNA-Seq to compare transcriptome responses of two Atlantic salmon populations to competition with ecological similar species with known dominance ranks. Overall, the effects of population on gene expression were higher than that of interspecific competition and there were both similarities and differences between Atlantic salmon populations in response to competition with ecologically similar species at the gene expression level.

Previous studies found that competition with rainbow trout or brown trout can have negative effects on growth and survival of Atlantic salmon, while competition with Chinook salmon has no negative effects (Van Zwol *et al.* 2012; Houde *et al.* 2015a). In this study, I found that Atlantic salmon had fewer responding genes to competition with Chinook salmon than to competition with rainbow trout or brown trout. This was expected because rainbow trout and brown trout are more aggressive than Atlantic salmon while Atlantic salmon is just as aggressive as Chinook salmon (Van Zwol *et al.* 2012; Houde *et al.* 2015a). I also found that the number of responding genes for Atlantic salmon in the presence of brown trout was similar (350 in LaHave and 384 in Sebago, respectively) between the two populations, which may indicate that brown trout stressed both Atlantic salmon populations similarly. However, the two Atlantic salmon populations showed substantial differences in the number of genes responding to competition with rainbow trout. That is, the number of genes responding to the presence of rainbow trout in the LaHave population was 3.6 times the number in the Sebago population. This suggests that the Sebago population may be more tolerant to the

presence of rainbow trout than the LaHave population, and thus the Sebago population may be more suitable for reintroduction in Lake Ontario as rainbow trout are common in the tributaries of the lake (Stanfield *et al.* 2006).

Among competition treatments within each population, LaHave Atlantic salmon reared with rainbow trout had the highest number of genes showing a significant response, while Sebago Atlantic salmon reared with brown trout had the most responding genes. Unlike previous results for the effects of interspecific competition on fitness-related traits which concluded that brown trout is the most serious competitor to Atlantic salmon and that rainbow trout can also have negative effects (Van Zwol *et al.* 2012; Houde *et al.* 2015a), my results implied that the most stressful competitor to the LaHave Atlantic salmon is rainbow trout while the most stressful competitor to the Sebago Atlantic salmon is brown trout. This implies that transcriptomic tools may be more sensitive to interspecies competition effects than commonly used fitness-related traits.

Although the functional categories of the genes showing a significant response to interspecific competition were broad and similar between the two populations, most of the genes responding to specific competitors were population-specific. Interestingly, most of the responding genes shared by the two Atlantic salmon populations in competition with Chinook salmon displayed the same gene expression regulation pattern, while most of the shared responding genes in competition with rainbow trout or brown trout showed contrasting gene expression regulation patterns. In particular, two somatostatin genes showed down-regulation in response to the presence of rainbow trout in both populations, and three somatostatin genes showed down-regulation in response to the presence of all three non-native salmonids in the Sebago population. Somatostatin is a hormone that

participates in multiple biological processes by inhibiting the release of pituitary hormones and gastrointestinal tract peptides (Burgus *et al.* 1973; Gahete *et al.* 2010). Somatostatin has been reported to regulate social behavior in cichlid fish (*Astatotilapia burtoni*) (Trainor & Hofmann 2006), with dominant males having larger somatostatin-containing neurons and higher expression of the *somatostatin* and *somatostatin receptor 3* genes in the hypothalamus relative to subdominant males (Hofmann & Fernald 2000; Trainor & Hofmann 2007). Additionally, Schunter *et al.* (2014) found the *somatostatin receptor 1* gene showed higher expression in the brain of territorial males than females in *Tripterygion delaisi* during the reproductive period. Although the functions of somatostatin genes in the spleen are not clear, the down-regulation of expression of these genes in competition with rainbow trout in both populations and in competition with the three species in the Sebago population may be adaptive because of the reported negative feedback regulation roles of somatostatin (Gahete *et al.* 2010).

## LIMITATIONS OF THIS STUDY

Although individuals may vary in stress response and tolerance, I pooled samples of four fish to minimize individual variation, and my aim was to examine the general influence of interspecific competition on the transcriptome of Atlantic salmon from two populations. Indeed, a statistical analysis of individuals within all treatments would have been more powerful, and may have detected a greater number of differentially expressed genes than my study. However, the general impacts of interspecific competition on the transcriptome in my study highlight a useful application of transcriptomic tools for source population selection for reintroduction and more broadly, conservation science.

## CONCLUSIONS

This study is the first report of transcriptomic responses to interspecific competition with ecologically similar species in fish. I found both similarities and differences in transcriptome responses to interspecific competition for the two Atlantic salmon populations. Overall, the Sebago population had fewer responding genes than the LaHave population, implying that the Sebago population was less affected by interspecific competition than the LaHave population, especially in competition with rainbow trout. This study can be added to the growing number of studies (Van Zwol *et al.* 2012; Houde *et al.* 2015a; b, 2016) indicating that the Sebago population likely has higher competitive ability than the LaHave population. Population differences in competitive ability can be reflected at the gene expression level and transcriptomic tools can be used to evaluate stress response and tolerance differences among populations for source population selection for conservation related applications. In particular, my transcriptome characterization highlights that organisms can possess different transcriptional responses to biotic stressors.

## REFERENCES

- Anders S, McCarthy DJ, Chen Y *et al.* (2013) Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nature Protocols*, **8**, 1765–1786.
- Anders S, Pyl PT, Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.
- Barton BA (2002) Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. *Integrative and Comparative Biology*, **42**, 517–525.
- Brennan RS, Galvez F, Whitehead A (2015) Reciprocal osmotic challenges reveal mechanisms of divergence in phenotypic plasticity in the killifish *Fundulus heteroclitus*. *The Journal of Experimental Biology*, **218**, 1212–1222.
- Buckland PR (2004) Allele-specific gene expression differences in humans. *Human Molecular Genetics*, **13**, R255–260.
- Burgus R, Ling N, Butcher M, Guillemin R (1973) Primary structure of somatostatin, a hypothalamic peptide that inhibits the secretion of pituitary growth hormone. *Proceedings of the National Academy of Sciences of the United States of America*, **70**, 684–688.
- Camacho C, Coulouris G, Avagyan V *et al.* (2009) BLAST+: architecture and applications. *BMC bioinformatics*, **10**, 421.
- Conesa A, Götz S, García-Gómez JM *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.
- Davis MA (2003) Biotic globalization: does competition from introduced species threaten biodiversity? *BioScience*, **53**, 481–489.
- Dimond P, Smitka J (2005) Evaluation of selected strains of Atlantic salmon as potential candidates for the restoration of Lake Ontario. *Trout Unlimited Canada Technical Report ON-012*, 41 pp.
- Fausch KD (2007) Introduction, establishment and effects of non-native salmonids: considering the risk of rainbow trout invasion in the United Kingdom. *Journal of Fish Biology*, **71**, 1–32.

- Feng J, Meyer CA, Wang Q *et al.* (2012) GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics*, **28**, 2782–2788.
- Gahete MD, Cordoba-Chacón J, Duran-Prado M *et al.* (2010) Somatostatin and its receptors from fish to mammals. *Annals of the New York Academy of Sciences*, **1200**, 43–52.
- Gilmour KM, DiBattista JD, Thomas JB (2005) Physiological causes and consequences of social status in salmonid fish. *Integrative and Comparative Biology*, **45**, 263–273.
- Hastings A (1987) Can competition be detected using species co-occurrence data? *Ecology*, **68**, 117–123.
- Hernandez ME, Martinez-Mota L, Salinas C *et al.* (2013) Chronic stress induces structural alterations in splenic lymphoid tissue that are associated with changes in corticosterone levels in wistar-kyoto rats. *BioMed Research International*, **2013**, 868742.
- Hodgins-Davis A, Townsend JP (2009) Evolving gene expression: from G to E to GxE. *Trends in Ecology & Evolution*, **24**, 649–658.
- Hofmann HA, Fernald RD (2000) Social status controls somatostatin neuron size and growth. *The Journal of Neuroscience*, **20**, 4740–4744.
- Houde ALS, Smith AD, Wilson CC, Peres-Neto PR, Neff BD (2016) Competitive effects between rainbow trout and Atlantic salmon in natural and artificial streams. *Ecology of Freshwater Fish*, **25**, 248–260.
- Houde ALS, Wilson CC, Neff BD (2015a) Competitive interactions among multiple non-native salmonids and two populations of Atlantic salmon. *Ecology of Freshwater Fish*, **24**, 44–55.
- Houde ALS, Wilson CC, Neff BD (2015b) Effects of competition with four nonnative salmonid species on Atlantic salmon from three populations. *Transactions of the American Fisheries Society*, **144**, 1081–1090.
- Jones ML, Stanfield LW (1993) Effects of exotic juvenile salmonines on growth and survival of juvenile Atlantic salmon (*Salmo salar*) in a Lake Ontario tributary. In: Production of Juvenile Atlantic Salmon, *Salmo salar*, in Natural Waters (eds Gibson RJ, Cutting RE), pp. 71-79. Canadian Special Publication of Fisheries and Aquatic Sciences.
- King RB, Ray JM, Stanford KM (2006) Gorging on gobies: beneficial effects of alien prey on a threatened vertebrate. *Canadian Journal of Zoology*, **84**, 108–115.
- Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.

- Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
- López-Maury L, Marguerat S, Bähler J (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nature Reviews Genetics*, **9**, 583–593.
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**.
- Manchester SJ, Bullock JM (2000) The impacts of non-native species on UK biodiversity and the effectiveness of control. *Journal of Applied Ecology*, **37**, 845–864.
- McDowall RM (2003) Impacts of introduced salmonids on native galaxiids in New Zealand upland streams: a new look at an old problem. *Transactions of the American Fisheries Society*, **132**, 229–238.
- Narum SR, Campbell NR (2015) Transcriptomic response to heat stress among ecologically divergent populations of redband trout. *BMC Genomics*, **16**.
- Olsvik PA, Lie KK, Jordal A-EO, Nilsen TO, Hordvik I (2005) Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. *BMC Molecular Biology*, **6**, 21.
- Peeler EJ, Oidtmann BC, Midtlyng PJ, Miossec L, Gozlan RE (2010) Non-native aquatic animals introductions have driven disease emergence in Europe. *Biological Invasions*, **13**, 1291–1303.
- Peters G, Schwarzer R (1985) Changes in hemopoietic tissue of rainbow trout under influence of stress. *Diseases of Aquatic Organisms*, **1**, 1–10.
- Petretto E, Mangion J, Dickens NJ *et al.* (2006) Heritability and tissue specificity of expression quantitative trait loci. *PLoS Genetics*, **2**, e172.
- Pimentel D, McNair S, Janecka J *et al.* (2001) Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, **84**, 1–20.
- Schulte PM (2004) Changes in gene expression as biochemical adaptations to environmental change: a tribute to Peter Hochachka. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, **139**, 519–529.
- Schunter C, Vollmer SV, Macpherson E, Pascual M (2014) Transcriptome analyses and differential gene expression in a non-model fish species with alternative mating tactics. *BMC Genomics*, **15**, 167.



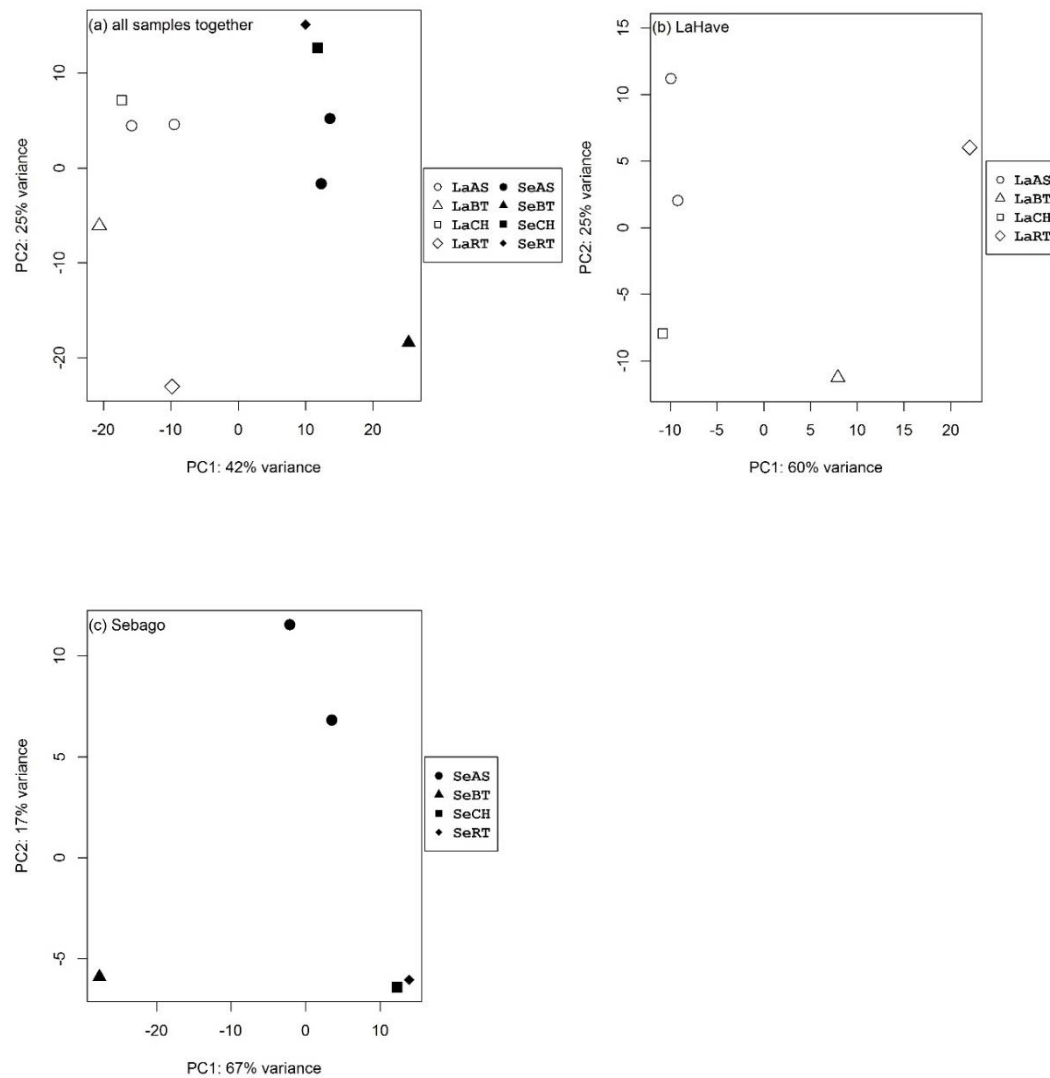
- Scott RJ, Noakes DLG, Beamish FWH, Carl LM (2003) Chinook salmon impede Atlantic salmon conservation in Lake Ontario. *Ecology of Freshwater Fish*, **12**, 66–73.
- Scott RJ, Poos MS, Noakes DLG, Beamish FWH (2005) Effects of exotic salmonids on juvenile Atlantic salmon behaviour. *Ecology of Freshwater Fish*, **14**, 283–288.
- Severns PM, Warren AD (2008) Selectively eliminating and conserving exotic plants to save an endangered butterfly from local extinction. *Animal Conservation*, **11**, 476–483.
- Sogge MK, Sferra SJ, Paxton EH (2008) Tamarix as Habitat for Birds: Implications for Riparian Restoration in the Southwestern United States. *Restoration Ecology*, **16**, 146–154.
- Stanfield LW, Gibson SF, Borwick JA (2006) Using a landscape approach to identify the distribution and density patterns of salmonids in Lake Ontario tributaries. *American Fisheries Society Symposium*, **48**, 601–621.
- Stewart TJ, Schaner T (2002) Lake Ontario salmonid introductions 1970 to 1999: stocking, fishery and fish community influences. In: *Lake Ontario Fish Communities and Fisheries: 2001 Annual Report of the Lake Ontario Management Unit*. Queen's Printer for Ontario, pp. 12.1-12.10.
- Tablado Z, Tella JL, Sánchez-Zapata JA, Hiraldo F (2010) The paradox of the long-term positive effects of a North American crayfish on a European community of predators. *Conservation Biology*, **24**, 1230–1238.
- Trainor BC, Hofmann HA (2006) Somatostatin regulates aggressive behavior in an African cichlid fish. *Endocrinology*, **147**, 5119–5125.
- Trainor BC, Hofmann HA (2007) Somatostatin and somatostatin receptor gene expression in dominant and subordinate males of an African cichlid fish. *Behavioural Brain Research*, **179**, 314–320.
- Trapnell C, Roberts A, Goff L *et al.* (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, **7**, 562–578.
- Turek KC, Pegg MA, Pope KL (2013) Review of the negative influences of non-native salmonids on native fish species. *Great Plains Research*, **23**, 11.
- Van Zwol JA, Neff BD, Wilson CC (2012) The effect of competition among three salmonids on dominance and growth during the juvenile life stage. *Ecology of Freshwater Fish*, **21**, 533–540.

Vitule JRS, Freire CA, Simberloff D (2009) Introduction of non-native freshwater fish can certainly be bad. *Fish and Fisheries*, **10**, 98–108.

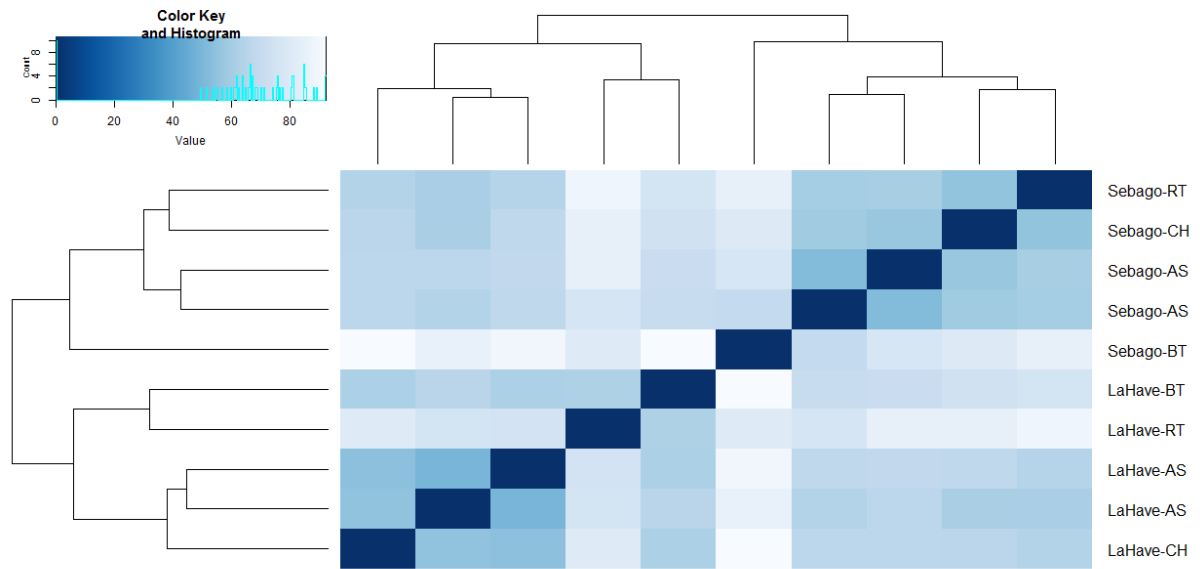
Wellband KW, Heath DD (2013) Environmental associations with gene transcription in Babine Lake rainbow trout: evidence for local adaptation. *Ecology and Evolution*, **3**, 1194–1208.

Whitehead A, Triant DA, Champlin D, Nacci D (2010) Comparative transcriptomics implicates mechanisms of evolved pollution tolerance in a killifish population. *Molecular Ecology*, **19**, 5186–5203.

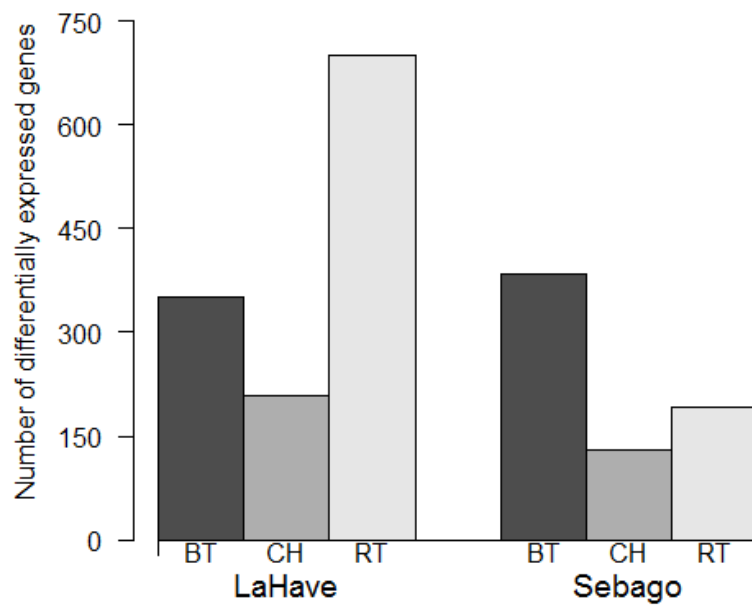
Ye J, Fang L, Zheng H *et al.* (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research*, **34**, W293-297.



**Figure 5.1** Principal component analysis based on expression levels for 2 000 selected genes which exhibited the highest expression variation among samples for (a) all the 10 samples, (b) the LaHave population samples ( $n = 5$ ) and (c) the Sebago population samples ( $n= 5$ ). Treatment symbols: LaAS indicates LaHave Atlantic salmon reared alone; LaBT, LaCH, and LaRT indicate LaHave Atlantic salmon reared with one of the three species: brown trout, Chinook salmon, and rainbow trout, respectively; SeAS indicates Sebago Atlantic salmon reared alone; SeBT, SeCH, and SeRT indicate Sebago Atlantic salmon reared with one of the three species: brown trout, Chinook salmon, and rainbow trout, respectively.



**Figure 5.2** Heatmap of sample-to-sample distances based on expression levels of all the genes. Treatment symbols: AS indicates Atlantic salmon reared alone; BT, CH, and RT indicate the Atlantic salmon population reared with one of the three species: brown trout, Chinook salmon, and rainbow trout, respectively.



**Figure 5.3** Number of responding genes of Atlantic salmon (*Salmo salar*) to the presence of brown trout (BT), Chinook salmon (CH), and rainbow trout (RT).

**Supplementary Table S5.1** Primers used for qRT-PCR.

Gene name	Forward primer	Reverse primer
60 kDa lysophospholipase	GAACATACGAGGCTACGAC	TCCATCAAACAGAGGCTAA
Alpha amylase	CTGGCTCCCAAAGGATACGC	CTCCACGGGCTGGTCAA
Aquaporin-1	ATCCTGGCTCAGATGCT	AACCCTTGTCCAACACTTATT
Carboxylic ester hydrolase	ACTGCCTCTACCTGAACATA	AGGGTCACCACGATTACA
Carboxypeptidase A1	ATCCTGGCTCATCCTGC	CCTTGGGTCCACGGTAA
Carboxypeptidase A2	TCAGGGTCAATGTGGAGTC	AGCGTACAGGCTGGAGTAG
Elastase-like serine protease	CCTATTGAGCCTCTGACCACC	TGTCTCCACTCACCGTCCC
Formin-binding protein 1	CGACGAATAAGTAACGAGG	GATGTAGTCTATCACAGTGCC
Hemoglobin subunit beta-1	CACTCCCGCAGCAATCAT	TTGTTGGCGTGGGTCTCG
Olfactomedin-4	CTTTGCTGCTGATGAGAC	GGTCAATGTAACGGGTG
Somatostatin-1A	TGCTCCAACGGTCACTCA	AGATCCACATCCTCCTGCT
Splicing factor 3B subunit 4-like	GGCTATGGCTTTGTTGAGTT	CCTTGTTGACACGGATGG
Trypsin 1	GCTGCTCACTGCTACAAG	AACTGCTCGCTACCCTCA
Trypsin 2	TGTGGAGGTGCGTCTGG	TTGAGGGTGGCGGGTTT
Ribosomal protein S20	CCCCTGTTGAGGCTGAG	TCCACGATAAGGTCTGC

**Supplementary Table S5.2** Mapping summary of the 10 samples to Atlantic salmon (*Salmo salar*) draft genome.

Sample name	Left/Input	Left/Mapped	Right/Input	Right/Mapped	Overall read mapping rate
LaAS1	15590578	83.60%	15590578	78.90%	81.30%
LaAS2	15580431	83.30%	15580431	78.90%	81.10%
LaCH	16726316	83.20%	16726316	79.00%	81.10%
LaBT	15832117	80.20%	15832117	72.40%	76.30%
LaRT	15406518	77.60%	15406518	68.20%	72.90%
SeAS1	16253754	82.00%	16253754	77.00%	79.50%
SeAS2	16941555	82.70%	16941555	78.10%	80.40%
SeCH	14921851	81.40%	14921851	76.60%	79.00%
SeBT	16245883	80.00%	16245883	75.30%	77.70%
SeRT	16431778	81.60%	16431778	75.50%	78.50%

**Supplementary Table S5.3** The expression of 10 genes which showed response to competition with all the three non-native salmonids for the LaHave Atlantic salmon (*Salmo salar*) population.

Gene ID	Description	RPKM			
		LaAS	LaBT	LaCH	LaRT
XLOC_006573	mesoderm induction early response 2 isoform X2	1.06	6.96	5.56	11.75
XLOC_007997	NA	1.80	6.35	0	9.36
XLOC_026040	PREDICTED: hypothetical protein LOC100636600	4.51	0	0	0
XLOC_035303	unnamed protein product	1.01	18.66	9.30	30.76
XLOC_038248	reverse transcriptase	0.75	0	2.88	4.79
XLOC_043264	fish virus induced TRIM	1.54	5.54	6.05	8.12
XLOC_044401	NA	3.32	0.04	0	0.08
XLOC_045484	NA	2.09	11.99	13.40	9.74
XLOC_051766	unnamed protein product	2.12	10.75	0	7.91
XLOC_059159	NA	1.60	7.50	20.15	8.16

Note: LaAS represents LaHave Atlantic salmon reared alone; LaBT, LaCH, and LaRT represent LaHave Atlantic salmon reared with brown trout, Chinook salmon, and rainbow trout, respectively.



**Supplementary Table S5.4** The expression of nine genes which showed response to competition with all the three non-native salmonids for the Sebago Atlantic salmon (*Salmo salar*) population.

Gene ID	Description	RPKM			
		SeAS	SeBT	SeCH	SeRT
XLOC_005337	Somatostatin-1 A precursor	16.53	1.32	0.12	2.06
XLOC_005338	Somatostatin-2 precursor	10.51	0.22	0.16	0.90
XLOC_005514	Somatostatin-1 A precursor	14.27	0.52	0.25	1.58
XLOC_010401	Dok-7-like isoform X1	1.09	4.65	0	4.98
XLOC_013971	phenylethanolamine N-methyltransferase-like	6.35	9.04	10.99	0.05
XLOC_025514	NA	11.93	0	0	0
XLOC_026425	glucagon-1 precursor	8.76	1.18	0	0.47
XLOC_031041	Insulin precursor	45.10	5.19	0.27	4.09
XLOC_035700	NA	24.35	32.63	1.06	0.37

Note: SeAS represents Sebago Atlantic salmon reared alone; SeBT, SeCH, and SeRT represent Sebago Atlantic salmon reared with brown trout, Chinook salmon, and rainbow trout, respectively.

**Supplementary Table S5.5** The expression of 23 genes which showed response to competition with Chinook salmon (*Oncorhynchus tshawytscha*) for both Atlantic salmon (*Salmo salar*) populations.

gene ID	Description	RPKM			
		LaAS	LaCH	SeAS	SeCH
XLOC_000035	hypothetical protein EAI_17313	1.56	0	2.17	0.24
XLOC_003299	apolipo A-II precursor	4.63	0.08	10.02	0.77
XLOC_012227	proglucagon II	0.48	3.82	2.18	0
XLOC_014649	Apolipo A-I precursor	1.23	0.03	1.97	0.16
XLOC_015883	AMBP precursor	1.52	0.16	3.38	0.29
XLOC_017724	apolipo A-I precursor	13.38	0.07	17.36	1.66
XLOC_020727	alpha-2-HS-glyco -like	2.31	0	4.05	0.62
XLOC_025691	beta-2-glyco 1-like	1.27	0	2.35	0.09
XLOC_026300	complement C5	1.63	0	2.38	0.15
XLOC_028644	Serpina1 , partial	6.22	0	8.91	1.05
XLOC_028948	apolipo B-100-like	0.99	0	2.61	0.22
XLOC_029296	fibrinogen gamma chain precursor	2.79	0.03	4.95	0.58
XLOC_029773	collagen alpha-1 chain-like isoform X3	0.60	3.89	6.56	0.60
XLOC_029784	Serotransferrin-1 precursor	3.54	0.46	7.73	1.49
XLOC_035374	Type-4 ice-structuring precursor	0.83	0	1.38	0.11
XLOC_038073	NA	1.11	4.55	2.86	9.81
XLOC_038626	apolipo A-II precursor	5.17	0	10.72	1.21
XLOC_039508	warm temperature acclimation-related 65 kDa	5.21	0.16	15.25	1.61
XLOC_040523	trout C-polysaccharide binding 1, isoform 1	4.11	0	9.37	0.58
XLOC_040524	trout C-polysaccharide binding 1, isoform 1	0.93	0	2.89	0.04
XLOC_043208	fibrinogen beta chain-like	3.78	0	5.63	0.72
XLOC_049651	serum albumin 2 precursor	7.45	0.29	14.16	2.39
XLOC_058750	NA	1.19	8.56	9.80	1.74

Note: LaAS represents LaHave Atlantic salmon reared alone; LaCH represents LaHave Atlantic salmon reared with Chinook salmon; SeAS represents Sebago Atlantic salmon reared alone; and SeCH represents Sebago Atlantic salmon reared with Chinook salmon.

**Supplementary Table S5.6** The expression of 13 genes which showed response to competition with brown trout (*Salmo trutta*) for both Atlantic salmon (*Salmo salar*) populations.

Gene ID	Description	RPKM			
		LaAS	LaBT	SeAS	SeBT
XLOC_004403	NA	1.26	6.69	2.40	0.14
XLOC_019565	NA	5.46	0	19.81	0
XLOC_023427	talin-1, partial	4.02	0.76	1.58	6.24
XLOC_034531	dnaJ homolog subfamily B member 5-like	1.31	0.18	0.08	0.68
XLOC_038248	reverse transcriptase	0.75	0	2.19	7.59
XLOC_039146	unnamed protein product	4.48	0.59	0.08	2.93
XLOC_042761	paternally-expressed gene 3 -like	0.48	3.87	0.07	1.54
XLOC_045381	hypothetical protein CAPTEDRAFT_85835, partial	2.30	0	0.76	3.92
XLOC_047231	hypothetical protein V500_07678	0.87	0	0.93	4.40
XLOC_050042	E3 ubiquitin- ligase HERC2 isoform X4	64.47	9.67	14.36	101.10
XLOC_054691	endonuclease domain-containing 1 -like	2.23	8.82	2.59	0.30
XLOC_059047	NA	0.30	9.31	15.46	1.88
XLOC_059836	NA	7.66	38.60	5.07	30.75

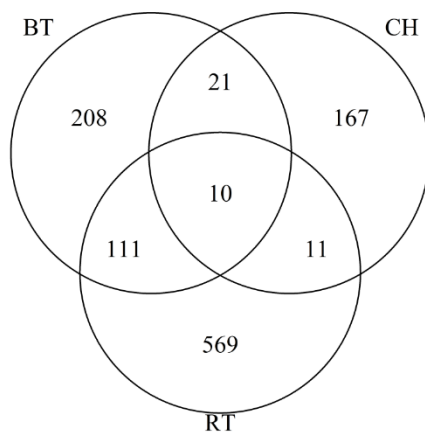
Note: LaAS represents LaHave Atlantic salmon reared alone; LaBT represents LaHave Atlantic salmon reared with brown trout; SeAS represents Sebago Atlantic salmon reared alone; and SeBT represents Sebago Atlantic salmon reared with brown trout.

**Supplementary Table S5.7** The expression of 19 genes which showed response to competition with rainbow trout (*Oncorhynchus mykiss*) for both Atlantic salmon (*Salmo salar*) populations.

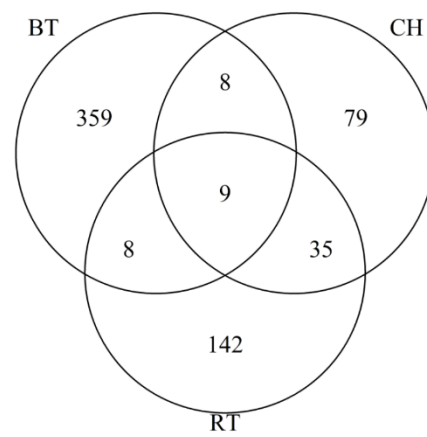
Gene ID	Description	RPKM			
		LaAS	LaRT	SeAS	SeRT
XLOC_002409	NA	14.78	1.17	0.41	6.85
XLOC_002993	NA	1.03	5.74	1.89	0.11
XLOC_004403	NA	1.26	9.28	2.40	0.20
XLOC_005337	Somatostatin-1A precursor	10.55	1.44	16.53	2.06
XLOC_005514	Somatostatin-1A precursor	6.87	0.52	14.27	1.58
XLOC_010417	transposase	0.04	1.07	1.73	0.26
XLOC_012848	RNA-directed DNA polymerase from mobile element jockey-like, partial	6.63	1.12	4.00	0
XLOC_024287	NA	0.03	1.28	0.45	0
XLOC_026425	glucagon-1 precursor	4.73	0.52	8.76	0.47
XLOC_032189	unnamed protein product	0.54	2.33	1.25	0.22
XLOC_034294	tyrosine aminotransferase	0.20	1.55	1.22	0
XLOC_038248	reverse transcriptase	0.75	4.79	2.19	9.24
XLOC_040524	trout C-polysaccharide binding 1, isoform 1	0.93	4.25	2.89	0.44
XLOC_042575	NA	1.13	8.75	3.89	0.45
XLOC_042576	NA	1.25	15.87	3.18	0
XLOC_048497	hypothetical protein VOLCADRAFT_70901	2.51	0.14	2.27	0.13
XLOC_051384	fatty acid synthase-like	3.06	0.40	0.44	2.55
XLOC_058750	NA	1.19	8.81	9.80	0.15
XLOC_060500	hypothetical protein	2.47	31.17	10.15	1.97

Note: LaAS represents LaHave Atlantic salmon reared alone; LaRT represents LaHave Atlantic salmon reared with rainbow trout; SeAS represents Sebago Atlantic salmon reared alone; and SeRT represents Sebago Atlantic salmon reared with rainbow trout.

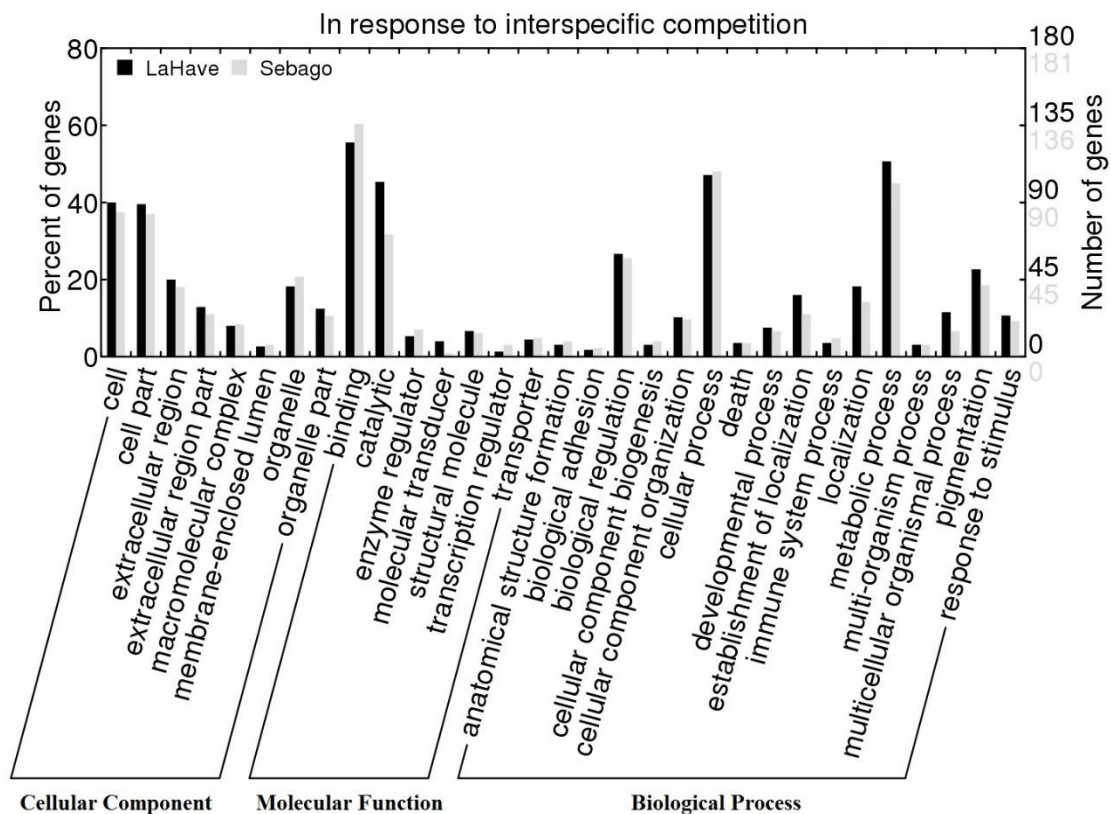
(a) LaHave



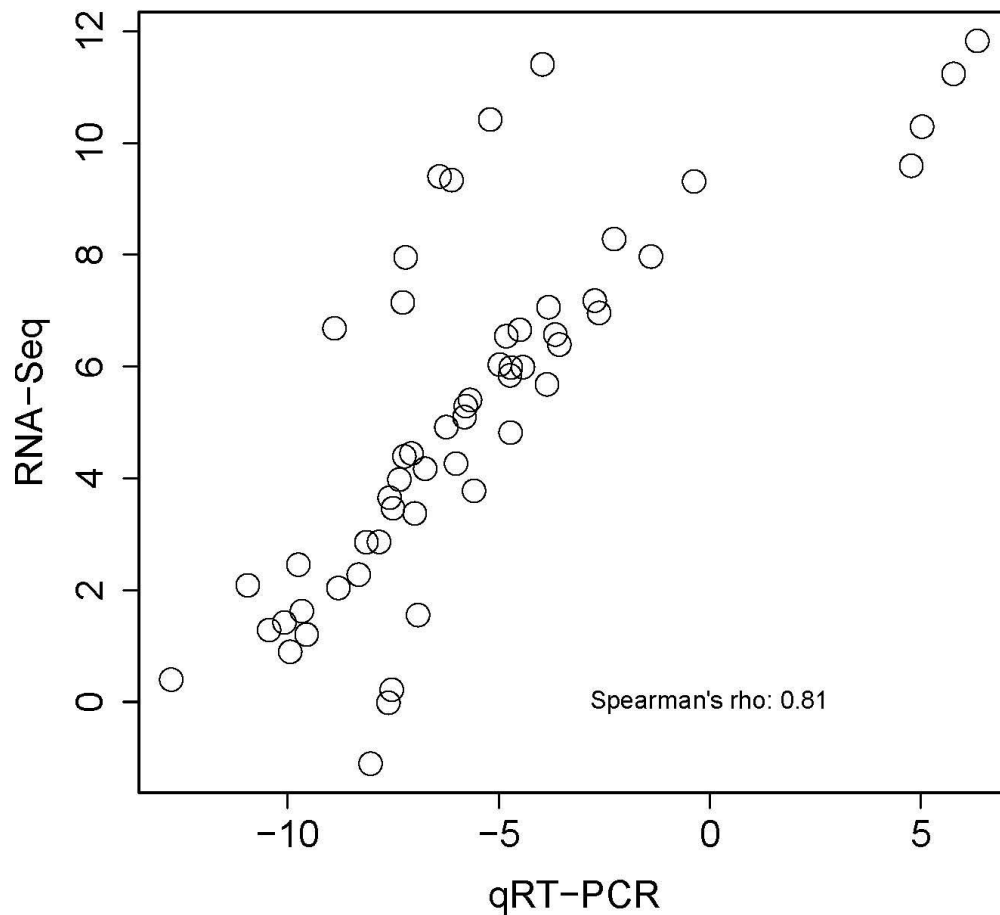
(b) Sebago



**Supplementary Figure S5.1** Venn diagram showing the overlap of differentially expressed genes in response to the presence of brown trout (BT), Chinook salmon (CH) and rainbow trout (RT) within the LaHave (a) and Sebago (b) Atlantic salmon populations.



**Supplementary Figure S5.2** Functional categorization of significantly different expressed genes in response to interspecific competition for the LaHave and Sebago Atlantic salmon (*Salmo Salar*) populations. For each population, responding genes were combined. GO terms containing less than five genes in both populations were not included.



**Supplementary Figure S5.3** Comparison of gene expression levels revealed by RNA-Seq and qRT-PCR for 14 genes. Expression of the 14 genes for RNA-Seq were  $\log_2$  transferred RPKM values from four samples: LaAS1 (LaHave Atlantic salmon reared alone sample 1), LaBT (LaHave Atlantic salmon reared with brown trout), SeAS1 (Sebago Atlantic salmon reared alone sample 1), and SeBT (Sebago Atlantic salmon reared with brown trout). qRT-PCR was used to quantify expression of the 14 genes for 12 individuals from the four treatments and then  $-\Delta C_T$  (calculated by  $\Delta C_{T_{\text{reference}}} - \Delta C_{T_{\text{targeted}}}$ ) values were averaged for each treatment. Spearman rank correlation analysis was conducted using the averaged  $-\Delta C_T$  for each treatment and  $\log_2$ RPKM.

## CHAPTER 6

### POPULATION-SPECIFIC RESPONSES TO INTERSPECIFIC COMPETITION IN THE GUT MICROBIOTA OF TWO ATLANTIC SALMON POPULATIONS<sup>5</sup>

#### INTRODUCTION

The intestine of normal animals harbors a great number and variety of bacteria which play an important role in animal health. Intestinal microbiota mediate a variety of biological processes and have been characterized as a “forgotten organ” (O’Hara & Shanahan 2006; Sommer & Bäckhed 2013). Symbiotic intestinal bacteria have long been recognized to aid in nutrient metabolism and absorption, and can provide vitamins to their host (Cummings & Macfarlane 1997; LeBlanc *et al.* 2013). Studies using germ-free animals demonstrated that intestinal bacteria are required for the differentiation of immune cells and normal development of the immune system and intestinal epithelium (Mazmanian *et al.* 2005; Olszak *et al.* 2012). Furthermore, gut microbiota play an important role in preventing colonization of opportunistic pathogens (Kamada *et al.* 2013). In addition, it has been reported that gut microbiota can regulate bone mass in mice and even host behavior in *Drosophila melanogaster* (Sharon *et al.* 2010; Sjögren *et al.* 2012). Clearly, changes in the composition and diversity of gut microbiota can affect the health of the host and the intestinal environment provided by the host can in turn affect the composition and dynamics of the gut microbial community as well.

---

<sup>5</sup> This is the outcome of joint research.



Although the community structure of the gut microbiota is the result of evolutionary interactions between the bacteria and their host (Ley *et al.* 2006), many external factors also can affect the composition and diversity of the gut microbiota. First, the physical environment experienced by the host (such as temperature and season) can have a great influence on the gut microbial community (Hagi *et al.* 2004; Neuman *et al.* 2016). Secondly, host physiological state can affect gut microbiota. For example, Bailey *et al.* (2010, 2011) found that stress exposure in mice significantly alters the relative abundances of certain types of gut bacteria and results in a greater incidence of colonization by pathogens. Furthermore, diet can impact the gut microbiota. Sullam *et al.* (2012) conducted a meta-analysis of 25 gut microbiota analyses in fish and found that the gut of herbivorous fish harbor more *Clostridiales*, *Bacteroidales*, and *Verrucomicrobiales* compared to omnivorous and carnivorous fish, while omnivorous fish harbor more *Rhizobiales*, *Fusobacteriales*, and *Planctomycetales* than carnivorous and herbivorous fish. Lastly, population source (captive or wild) and geographic variation also can affect composition and diversity of gut microbiota (Linnenbrink *et al.* 2013; Kreisinger *et al.* 2014; Stevenson *et al.* 2014; Zarkasi *et al.* 2014). The effects of those factors on gut microbiota imply that ecological challenges and environmental stresses organisms encounter can indirectly or directly affect their gut microbiota.

Differences in response to, and tolerance of, environmental stress among populations of the same species have been reported in a variety of fish species (DeKoning *et al.* 2004; Fangue *et al.* 2006; Whitehead *et al.* 2010; Côte *et al.* 2012). Although population differences in stress response have been well characterized for physiological and life history traits, population-level differences in how fish gut microbiota respond to

ecological stress have not been explored. As the gut microbiota is clearly a critical component of fish health, the response of the microbiota to host stress levels is an obvious factor to examine when selecting source populations for aquaculture, restocking, and reintroduction. Given the close relationship between host physiology, health, and gut microbiota, demonstrating stress response differences in gut microbiota composition among populations in common garden experiments will provide insight into predicting population performance differences under the stressful conditions associated with a novel environment.

Atlantic salmon (*Salmo salar*) was once a native species in Lake Ontario, but was extirpated late in the 19<sup>th</sup> century mainly as a result of habitat fragmentation and degradation (Crawford 2001). Due to its important economic, cultural and ecological roles, there have been increasing efforts to reintroduce Atlantic salmon into Lake Ontario for over 30 years; however, those reintroductions have been unsuccessful. After Atlantic salmon was extirpated, Chinook salmon *Oncorhynchus tshawytscha*, coho salmon *O. kisutch*, rainbow trout *O. mykiss*, and brown trout *S. trutta* were successfully introduced into Lake Ontario and its tributaries to provide recreational fishing opportunities (Stewart & Schaner 2002). The establishment of those four non-native salmonids has been proposed as a significant barrier to the successful reintroduction of Atlantic salmon into Lake Ontario because of intense interspecific competition at both the juvenile and adult stages due to niche overlap, impeding the successful reintroduction of Atlantic salmon (Scott *et al.* 2003, 2005; Van Zwol *et al.* 2012; Houde *et al.* 2015a, b, 2016). Although stressful interspecific competitive interactions have been shown to affect growth and survival (Houde *et al.* 2015a, b), the mechanisms behind those effects are not clear, as

stress can affect many aspects of organisms at different levels. One known outcome of stress is a detrimental change in the intestinal microbial community that impacts the host through multiple pathways of bidirectional interaction between gut microbiota and their host (Carabotti *et al.* 2015). However, the effect of interspecific competition on the gut microbiota has not been explored in any species.

To explore the role of interspecific competition on the gut microbiota and to test for evidence for competition stress tolerance differences between source populations for Atlantic salmon reintroduction, I conducted interspecific competition experiments for two Atlantic salmon populations exposed to the four established non-native salmonids of Lake Ontario. I used next generation sequencing of the *16S rRNA* gene to characterize the composition and diversity of intestinal microbiota of juvenile Atlantic salmon. I aimed to test three hypotheses in this study. First, I hypothesize that there would be microbial community differences between the two source populations in diversity and composition, and these differences would reflect co-evolutionary differences in host and microbiota dynamics of the two populations. Second, I hypothesize that there would be greater microbial community response to interspecific competition in the LaHave population relative to the Sebago population, reflecting the reported pattern of lower interspecific competition tolerance in the LaHave population relative to the Sebago population (Houde *et al.* 2015a, b, 2016). Last, as chronic stress is known to reduce immunity and disease resistance in fish (Barton 2002), I hypothesize that interspecific competition will result in decreased relative abundance of beneficial bacteria and increased relative abundance of opportunistic pathogens within the gut in both populations, but to a lesser extent in the Sebago population. My results have important implications for the understanding of the

nature of the co-evolution of the host and their microbiota. Furthermore, my work shows that the interrelationship between the host and their gut microbiota is a critical factor to consider when selecting source populations for the conservation and management of species at risk and commercially exploited species.

## MATERIALS AND METHODS

### *Interspecific competition and sample collection*

Currently, two Atlantic salmon populations are being used for reintroduction into Lake Ontario: LaHave and Sebago. The LaHave is an anadromous population which originates from LaHave River, Nova Scotia, whereas the Sebago is a landlocked population from Sebago Lake, Maine. For competing species, I used four non-native salmonids (Chinook salmon, coho salmon, rainbow trout, and brown trout) which have been introduced and are established in Lake Ontario tributaries. Juveniles (fertilized in November 2011) of all the five species were provided by the Ontario Ministry of Natural Resources and Forestry (OMNRF), Canada. Details about the fish breeding are provided in Houde *et al.* (2015b).

In September 2012, the Atlantic salmon and competing species were transferred to artificial stream tanks, commencing the interspecific competition experiment. Each artificial stream tank included a riffle and a pool microhabitats (160 cm long for the riffle and 80 cm long for the pool). Details about the artificial stream tank design are provided in Houde *et al.* (2015b). There were six treatments for each Atlantic salmon population: Atlantic salmon reared alone (32 Atlantic salmon), Atlantic salmon reared in a 1:1 ratio with one of the four non-native species (16 Atlantic salmon and 16 one of the non-native

species), and Atlantic salmon reared with all the four species combined (16 Atlantic salmon, four Chinook salmon, four coho salmon, four rainbow trout, and four brown trout). Each trial was replicated. To minimize any differences in performance caused by genetic effects, Atlantic salmon from each of the two populations were comprised of equal numbers of fish from eight full-sib families in each tank. The fish were fed commercial pellet feed at 1% of their body mass per day from January to April and 3% of their body mass per day in other months. Previous studies demonstrated that juvenile Atlantic salmon, rainbow trout and brown trout prefer riffle microhabitats whereas Chinook salmon and coho salmon prefer pool microhabitats (Hartman 1965; Morantz *et al.* 1987; Holecek *et al.* 2009). Therefore, I expected that the five treatments should result in a range of competitive effects on the Atlantic salmon gut microbiota when they competed for feed and microhabitat.

At the end of July 2013, after 10 months in the artificial stream tanks, six to nine Atlantic salmon from each tank were randomly collected and humanely euthanized by an overdose of buffered MS-222. I collected intestinal content of Atlantic salmon from both replicated tanks for each treatment, except for the LaHave population reared with brown trout because there were only surviving Atlantic salmon in one replicate tank. Prior to dissection, the fish were externally disinfected using 75% ethanol and subsequently opened using a sterile scalpel. Intestinal contents were collected and stored at -20 °C immediately, and were transferred to the lab later on ice. In addition, I collected 500 mL water samples from four tanks for microbial analysis. The water was filtered using Supor®200 Membrane Filter with 0.2 µm pore size (Pall Corporation, Mississauga, ON, Canada) and the filter was stored frozen for later DNA extraction.

### *DNA extraction, PCR, and library preparation*

Bacterial genomic DNA was extracted using the E.Z.N.A.® Stool DNA Kit (Omega Bio-tek, Norcross, GA, USA) following the supplier's instructions. In total, I extracted bacterial DNA from the intestinal content of 178 fish and the four water samples. The V5 and V6 regions of the bacterial *16S rRNA* gene were PCR amplified using previously reported primers (Sogin *et al.* 2006; Roesch *et al.* 2007). I used two rounds of PCR to first amplify the target region and then to ligate adaptor and barcode sequences for next generation sequencing. To reduce the incidence of PCR artifacts in the first round, I used the minimum number of PCR cycles such that enough target DNA was amplified to show faint bands on agarose gel (Lenz & Becker 2008). The 1<sup>st</sup> PCR was conducted in 25 µL reactions consisting of 2.5 µL 10 × Buffer (including Mg<sup>2+</sup>), 0.5 µL 10 mM dNTP, 0.4 µL 10 µM forward primer (V5F, Table 6.1), 0.4 µL 10 µM reverse primer (V6R, Table 6.1), 0.25 µL BSA, 1 Unit Taq and 1 µL DNA. The thermal cycle protocol for the first PCR was: 95 °C for 150 s followed by 26 cycles of 95 °C for 30 s, 53 °C for 30 s, and 72 °C for 1 min, followed by a final elongation at 72 °C for 10 min. The PCR product was checked on agarose gel and then purified using Agencourt AMPure XP beads (Beckman Coulter Genomics GmbH, Mississauga, ON, Canada). The second PCR (to ligate the adaptor and barcode sequences) was conducted in 25 µL reactions consisting of 2.5 µL 10 × Buffer (including Mg<sup>2+</sup>), 0.5 µL 10 mM dNTP, 0.4 µL 10 µM forward primer (UniA, Table 6.1), 0.4 µL 10 µM reverse primer (UniB, Table 6.1), 0.25 µL BSA, and 15 µL of the purified first PCR product. The protocol of the second PCR was 95 °C for 150 s, then 7 cycles of 95 °C for 30 s, 60 °C for 30 s, and 72

°C for 1 min, followed by a final elongation at 72 °C for 5 min. The forward and reverse primers used in the first PCR had a 12bp tail at the 5' end which complemented the 3' end of the corresponding primer used in the second PCR (Table 6.1). The forward primer in the second PCR included individually unique 10 - 12 bp barcode sequences that allowed me to sort final sequence reads to the original sample after multiplexed sequencing (Table 6.1). The second PCR products from all the samples were mixed together and purified using QIAquick Gel Extraction Kit (QIAGEN, Toronto, ON, Canada). The purified PCR product mix was then run on an Agilent 2100 Bioanalyzer with a High Sensitive DNA chip (Agilent Technologies, Mississauga, ON, Canada) to measure the DNA concentration. The library was then diluted to 26 pmol/L. The sequencing reaction was run on an Ion PGM™ System using the Ion PGM™ Sequencing 400 Kit and an Ion 318™ Chip (Thermo Fisher Scientific, Burlington, ON, Canada).

#### *Bioinformatic and statistical analysis*

Bioinformatic analyses were conducted using Quantitative Insights Into Microbial Ecology (QIIME) 1.8 (Caporaso *et al.* 2010). After de-multiplexing and quality filtering of the raw sequence reads, reference-based and *de novo* chimeras were removed from the cleaned sequences and Operational Taxonomic Unit (OTU) clustering was performed with a 0.97 threshold using urearch (Edgar 2010). The representative sequence for each OTU was selected using the most abundant method for assigning taxonomy using RDP Classifier program with a minimum 80% confidence level (Wang *et al.* 2007). As my focus is on the functional significance of changes in the gut bacterial community, the unclassified sequences at the domain level and sequences belonging to the *Archaea*

domain and *Cyanobacteria* phylum were removed from the OTU table (Wong *et al.* 2013).

Four alpha diversity metrics (chao1, shannon index, observed species number, and phylogenetic distance) for each sample were estimated using QIIME. I applied a linear mixed effects model to test the effects of population and competition treatment on alpha diversity indices which were calculated based on 2 015 sequences per sample. In the linear mixed effects model, the effects of population, treatment, and the interaction between population and treatment were fixed effects and the replicate tank effect was a random effect.

To test population and treatment effects on community divergence (beta diversity), the OTUs for each sample were rarefied to 2 000 sequence/sample and the weighted UniFrac distance matrix was computed (Lozupone & Knight 2005). Then, *adonis* analyses were performed to test for the effects of population, treatment, and the interaction between population and treatment on this distance matrix using the *vegan* R package (Oksanen *et al.* 2015). To analyze population-specific treatment effects on gut microbiota, I computed weighted UniFrac distance within each population separately and then conducted *adonis* analysis in the two populations.

To study the effects of population and interspecific competition on gut microbiota at the individual OTU level, I analyzed relative abundance for the 180 most abundant OTUs that appeared in at least 70% of the gut samples. To test for differences of relative OTU abundance between the two populations, I applied the Welch's t-test and *P* values were corrected for multiple simultaneous comparisons using Benjamini-Hochberg FDR in Statistical Analysis of Metagenomic Profiles (STAMP) v2.0.8 (Parks *et al.* 2014). To



test for interspecific competition effects on relative OTU abundance among the competition treatments within each population, I conducted a one-way ANOVA followed by a Tukey-Kramer post-hoc test for the two populations separately using STAMP v2.0.8 (Parks *et al.* 2014).

To test for treatment effects on the relative abundance of beneficial bacteria in the gut, I applied a one-way ANOVA within each population to analyze the relative abundance of the *Bacillus* genus and seven lactic acid bacteria genera (*Carnobacterium*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus* and *Streptococcus*), because many species from those genera have been proposed as probiotics and have documented positive effects in aquaculture applications (Merrifield *et al.* 2010; Merrifield & Carnevali 2014). To test for the effect of the interspecific competition on the relative abundance of potential pathogens for each population, I used a one-way ANOVA for relative abundance in four genera (*Aeromonas*, *Flavobacterium*, *Mycobacterium* and *Vibrio*), because some species from those genera are known fish pathogens (Bøgwald & Dalmo 2014; Miller *et al.* 2014).

## RESULTS

### *Summary of sequencing and core OTUs*

The raw sequences generated in this study have been submitted to NCBI Sequence Read Archive (SRA) database (Accession number: SRP071211). After de-multiplexing and filtering out poor quality sequences, I obtained 4 111 310 high quality sequences. The number of sequence per sample ranged from 2 220 to 54 951 with an average of 22 590 (Supplementary Figure S6.1). The average numbers of sequence reads

per sample in each of the two Atlantic salmon populations were very similar (Supplementary Figure S6.2). In total, 3 978 bacterial OTUs were identified. While the definition of core OTU varies among studies, I define core OTU as the OTUs which are present in 70% of gut samples. Among all the fish gut samples, I found 180 core OTUs and those OTUs accounted for 74.7% to 90.1% of sequences in each treatment (Supplementary Figure S6.3).

#### *Bacterial community composition*

I identified 26 bacterial phyla across the two Atlantic salmon populations, and 14 phyla in the four water samples. Among the ten most abundant phyla (*Proteobacteria*, *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, *Planctomycetes*, *Chlamydiae*, *Verrucomicrobia*, *Chloroflexi*, *TM7* and *Fusobacteria*), *Proteobacteria* and *Bacteroidetes* showed higher relative abundance in the water than in the gut microbial communities, while the other eight phyla were at higher relative abundance in gut microbiota (Supplementary Figure S6.4).

Within the gut microbiota of the two Atlantic salmon populations, *Proteobacteria* (64.3%-83.6%) was the most common phylum followed by *Firmicutes* (12.9%-23.2%; Supplementary Figure S6.5). At the lower taxonomic levels, there was substantially more variation among the treatments within populations. At the class level, *Gammaproteobacteria* was the most common, while the second most abundant bacterial class varied from treatment to treatment: *Betaproteobacteria* was the second most abundant bacterial class for most treatments; *Bacilli* and *Clostridia* were the second most abundant bacterial classes for two and three treatment groups respectively (Figure 6.1).

At the genus level, about half the reads could not be assigned to a single genus; however, the dominant genus was *Pseudomonas* in all the treatment groups (Supplementary Figure S6.5). *Acinetobacter*, *Deefgea*, *Rhodobacter*, *Flavobacterium* and *Lactobacillus* also showed high abundance (Supplementary Figure S6.6).

#### *Effects of population and treatment on bacterial diversity*

All four alpha diversity estimates exhibited significant differences between the two populations, while the effects of treatment, and the interaction between population and treatment were not significant. The Sebago population had significantly higher alpha diversity than the LaHave population for all four metrics. The *adonis* analysis of beta diversity based on all the fish gut microbiota samples showed that population and treatment had significant effects on the weighted UniFrac distance ( $R^2 = 0.083$ ,  $P = 0.001$ ;  $R^2 = 0.054$ ,  $P = 0.009$ , respectively). The effect of interaction between population and treatment on weighted UniFrac distance was not significant. When the weighted UniFrac distance matrices were computed for each population separately, treatment only showed as a significant effect in the LaHave population ( $R^2 = 0.122$ ,  $P = 0.011$ ).

#### *Population and treatment effects at the OTU level*

Welch's t-test on the 180 OTUs showed that 27 OTUs had significantly different relative abundances between the two populations across treatments after FDR correction (Figure 6.2). Of the 27 OTUs, 10 OTUs showed higher relative abundance in the LaHave population, and all of those belong to two families: *Aeromonadaceae* and *Shewanellaceae* (Figure 6.2). Seventeen OTUs showed higher relative abundance in the

Sebago population relative to the LaHave population, and 10 of those belong to the *Rhodobacteraceae* family (Figure 6.2).

For the effect of interspecific competition on the composition and diversity of gut bacterial communities within each population, there were 13 OTUs that showed significant differences among treatments within the LaHave population (Figure 6.3). For all 13 OTUs, the gut microbiota from Atlantic salmon reared alone and Atlantic salmon reared with Chinook salmon showed similar relative abundances while the other four treatments showed lower abundances (Figure 6.3). Among the 13 OTUs, seven OTUs belong to the *Flavobacteriales* order, five OTUs belong to the *Lactobacillales* order and one OTU belongs to the *Enterobacteriales* order. There were no OTUs which showed significant differences in relative abundance among treatments in the Sebago population.

#### *Differences in beneficial bacteria and opportunistic pathogens*

Six lactic acid bacteria genera showed significant differences among treatments within the LaHave population (Figure 6.4). The lactic acid bacteria genera showed similar relative abundance in the gut of Atlantic salmon reared alone and Atlantic salmon reared with Chinook salmon, and showed reduced relative abundance in the other four treatments (Figure 6.4). Within the Sebago population, the lactic acid bacteria genera showed no significant difference among treatments. The *Bacillus* genus showed no significant difference among treatments in either population.

For the potential pathogens, no *Aeromonas* genus was detected in the gut contents and there was no significant difference in the relative abundance of the combined OTUs in the *Flavobacterium*, *Mycobacterium* and *Vibrio* genera among treatments within each

population (Supplementary Figure S6.6). While I did observe a significant effect of interspecific competition on a single OTU within the *Flavobacterium* genus (OTU 39; Figure 6.3b) in the LaHave population, the response pattern across this genus was not significant (Supplementary Figure S6.7).

## DISCUSSION

The composition and diversity of gut microbiota is known to be determined by a combination of genetic and environmental factors (McKnite *et al.* 2012; Sullam *et al.* 2012; Parks *et al.* 2013; Stevenson *et al.* 2014; Yun *et al.* 2014). Although reared under the same conditions, I found significant differences between the two study populations using different measures of microbial community composition. Those differences are likely due to different evolutionary histories which shaped the co-evolution of the host and their gut microbiota. Of the 10 OTUs that showed higher relative abundance in the LaHave population, seven belong to the *Aeromonadacea* family. The *Aeromonas* genus of the *Aeromonadacea* family contains two important fish pathogens: *Aeromonas hydrophila* and *Aeromonas salmonicida* which infect various fish species (Ringø *et al.* 2010). Seven of the 17 OTUs that showed higher abundance in the Sebago population belong to the genus *Rhodobacter* and two members of the *Rhodobacter* genus are used as probiotics in aquaculture in China (Qi *et al.* 2009). Thus the observed population differences are consistent with the Sebago population harboring higher abundances of beneficial bacteria and lower abundances of opportunistic pathogens relative to the LaHave population. This effect is despite the two groups having been reared in a common environment since fertilization. Although the effects of gut microbiota on the

host are complex, this pattern of gut microbiotic differences indicates that the Sebago population has advantageous gut microbiota relative to the LaHave population across the competition treatments and may thus reflect a higher interspecific competition tolerance in the Sebago population (Houde *et al.* 2015b, 2016).

I found that the Sebago population had significantly higher alpha diversity than the LaHave population. It has been reported that stress can reduce the alpha diversity of gut microbiota and change the relative composition of bacteria in mice (Bailey *et al.* 2010, 2011). Although I did not detect significant interspecific competition treatment effects on alpha diversity in Atlantic salmon, the population-level differences in diversity may reflect stress effects across all competition treatments combined with rearing stress. This is supported by previous studies that demonstrated that the Sebago population has higher competitive ability and are less affected by interspecific competition than the LaHave population (Houde *et al.* 2015b, 2016). As these fish were reared in a common environment and provided the same feed, these population-level difference support previous reported genetics effect on gut microbiota (Goodrich *et al.* 2014). The fundamental differences in gut microbiota between the two populations implied that the two populations may retain their ancestral co-evolved microbial community despite years of rearing in a common artificial environment or have experienced different co-evolutionary pressures in the hatchery environment.

Houde *et al.* (2015a, b, 2016) reported strong interspecific competition effects on the growth and survival of Atlantic salmon, which was likely a reflection of tertiary responses to stress caused by the interspecific competition (Barton 2002). In my study, I found significant interspecific competition effects on the abundance of specific OTUs

and on the weighted UniFrac distance, but only in one of the study populations. Among the 13 OTUs that showed significant competition treatment effects within the LaHave population, all showed no competition effect with Chinook salmon. This outcome was expected because previous studies showed that Chinook salmon had no negative effects on Atlantic salmon when they were reared together (Houde *et al.* 2015a, b). However, those same studies reported that interspecific competition with coho salmon had no negative effects on growth and survival (Houde *et al.* 2015a, b), but I found that competition with coho salmon resulted in patterns of altered OTU abundance similar to those in the Atlantic salmon reared with brown trout and rainbow trout. This indicates that the gut microbiota is more sensitive than growth and survival traits to stress related to interspecific competition. Curiously, I did not find any OTUs with significantly higher abundance in response to competition with the more aggressively competitive species. This may be due to high OTU abundance variation among individuals that responded to interspecific competition with elevated OTU abundance; that is, I had low power to detect those changes as statistically significant. Differences in the 13 OTUs among treatments indicate gut microbiota showed response to interspecific competition and the response in gut microbiota is population specific. The high sensitivity of gut microbiota to stress demonstrates the potential of gut microbiota as a biomarker to evaluate stress response and tolerance differences among individuals and populations.

Gut microbiota plays an important role in the health of their host and the gut contains both beneficial bacteria and opportunistic pathogens (Kamada *et al.* 2013). While chronic stress is known to have detrimental effect on organism's health and disease resistance (Barton 2002; Sommer & Bäckhed 2013), the mechanism is not well

understood. One possible mechanism explaining how that interspecific competition can negatively affect the growth and survival of Atlantic salmon is through changes in the beneficial bacteria and potential pathogen in the gut. In this study, I found that interspecific competition has profound impacts on the abundance of 13 OTUs in one of the two Atlantic salmon populations; however, the functional significance of those changes are not obvious. My analysis of known or suspected beneficial and pathogenic bacteria was designed to address the functional component of gut microbiotic response in Atlantic salmon. I found that interspecific competition with the more aggressive competitors decreased the relative abundance of beneficial lactic acid bacteria in the LaHave, but not in the Sebago population. Lactic acid bacteria are generally considered beneficial because they not only enhance immune response and positively affect immune systems of the host (Perdigón *et al.* 2001), but they also function in preventing the colonization of pathogens, possibly by producing bacteriocin or competing with pathogens for nutrients (Ringø 2008). Many species of lactic acid bacteria have been used widely as probiotics to increase growth and disease resistance in fishes, including salmonids (Merrifield *et al.* 2010; Merrifield & Carnevali 2014). The reduction in lactic acid bacteria in Atlantic salmon reared with coho salmon, rainbow trout, and brown trout indicate that interspecific competition can cause a loss of beneficial gut microbiota. The likely mechanism for this loss of probiotic bacteria in the LaHave Atlantic salmon is competition-related stress as previous studies showed that stress decreases the abundance of *Lactobacillus* in human and monkey gut microbiota (Bailey *et al.* 2004; Knowles *et al.* 2008).



Given the profound impact of interspecific competition on the gut microbiota in the LaHave Atlantic salmon, it was surprising that I did not detect any significant increase in the selected pathogenic genera (*Flavobacterium*, *Mycobacterium* and *Vibrio*). Bailey *et al.* (2010) reported elevated abundance of *Citrobacter rodentium* in the gut of mice subjected to prolonged restraint stress after they were challenged by this pathogen via oral gavage. In my study, the lack of obvious pathogenic microbial response to the treatments may be due to the fish being reared in a hatchery such that the opportunity for pathogenic colonization was at such low levels they did not occur at levels sufficient to detect. While it is possible that the competition stress in my experiment affects gut microbiota composition, but not the host's susceptibility to bacterial disease; it is not clear how such an anomalous response could be mediated.

The differences between the two study populations (LaHave and Sebago) in the general composition of their gut microbiota and in their specific response to interspecific competition can only be explained by different host-microbiota co-evolutionary histories, despite a few generations of hatchery rearing (three generations for the LaHave population and one generation for the Sebago population). Although horizontal transfer of bacteria between Atlantic salmon and non-native salmonids is possible, no study has shown such an effect and it is likely a small effect if it is present. My results have important applications in source population selection for reintroduction. The higher gut microbiota alpha diversity, higher abundance of beneficial bacteria, and lower microbial community change in response to interspecific competition indicate that the Sebago population is more suitable than the LaHave population for reintroduction into Lake Ontario. The juvenile Sebago Atlantic salmon are more tolerant to stress caused by

competition with the non-native salmonids known to be present in the tributaries of Lake Ontario. More generally, my results also indicate that gut microbiota is a good candidate as a biomarker for stress tolerance and thus for the selection of source populations for reintroduction, conservation, aquaculture, and other applications.

## CONCLUSIONS

This is the first study to report the effects of interspecific competition on gut microbial communities. I detected significant differences in gut microbiota and alpha diversity between the two Atlantic salmon populations as well as profound differences in their response to interspecific competition. My study shows that population differences and population-specific responses in gut microbiota can be part of the mechanism involved in differential performance under competition with ecologically similar species. The characterization of changes in gut microbiota is now possible for ecological and evolutionary studies of competition and co-evolution and my work highlights the dynamic role of the host's gut microbiota in both evolutionary and ecological processes.

## REFERENCES

- Bailey MT, Dowd SE, Galley JD *et al.* (2011) Exposure to a social stressor alters the structure of the intestinal microbiota: implications for stressor-induced immunomodulation. *Brain, Behavior, and Immunity*, **25**, 397–407.
- Bailey MT, Dowd SE, Parry NMA *et al.* (2010) Stressor exposure disrupts commensal microbial populations in the intestines and leads to increased colonization by *Citrobacter rodentium*. *Infection and Immunity*, **78**, 1509–1519.
- Bailey MT, Lubach GR, Coe CL (2004) Prenatal stress alters bacterial colonization of the gut in infant monkeys. *Journal of Pediatric Gastroenterology and Nutrition*, **38**, 414–421.
- Barton BA (2002) Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. *Integrative and Comparative Biology*, **42**, 517–525.
- Bøggwald J, Dalmo RA (2014) Gastrointestinal Pathogenesis in Aquatic Animals. In: *Aquaculture Nutrition* (eds Merrifield D, Ringø E), pp. 53–74. John Wiley & Sons Ltd.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, **7**, 335–336.
- Carabotti M, Scirocco A, Maselli MA, Severi C (2015) The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Annals of Gastroenterology*, **28**, 203–209.
- Côte J, Roussel JM, Le Cam S, Bal G, Evanno G (2012) Population differences in response to hypoxic stress in Atlantic salmon. *Journal of Evolutionary Biology*, **25**, 2596–2606.
- Crawford SS (2001) Salmonine introductions to the Laurentian Great Lakes: an historical review and evaluation of ecological effects. *Canadian Special Publication of Fisheries and Aquatic Sciences* 132.
- Cummings JH, Macfarlane GT (1997) Role of intestinal bacteria in nutrient metabolism. *Clinical Nutrition*, **16**, 3–11.
- DeKoning ABL, Picard DJ, Bond SR, Schulte PM (2004) Stress and interpopulation variation in glycolytic enzyme activity and expression in a teleost fish *Fundulus heteroclitus*. *Physiological and biochemical zoology*, **77**, 18–26.

- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.
- Fangue NA, Hofmeister M, Schulte PM (2006) Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *The Journal of Experimental Biology*, **209**, 2859–2872.
- Goodrich JK, Waters JL, Poole AC *et al.* (2014) Human genetics shape the gut microbiome. *Cell*, **159**, 789–799.
- Hagi T, Tanaka D, Iwamura Y, Hoshino T (2004) Diversity and seasonal changes in lactic acid bacteria in the intestinal tract of cultured freshwater fish. *Aquaculture*, **234**, 335–346.
- Hartman GF (1965) The role of behavior in the ecology and interaction of underyearling coho salmon (*Oncorhynchus kisutch*) and steelhead trout (*Salmo gairdneri*). *Journal of the Fisheries Research Board of Canada*, **22**, 1035–1081.
- Holecek DE, Cromwell KJ, Kennedy BP (2009) Juvenile Chinook salmon summer microhabitat availability, use, and selection in a central Idaho Wilderness stream. *Transactions of the American Fisheries Society*, **138**, 633–644.
- Houde ALS, Smith AD, Wilson CC, Peres-Neto PR, Neff BD (2016) Competitive effects between rainbow trout and Atlantic salmon in natural and artificial streams. *Ecology of Freshwater Fish*, **25**, 248–260.
- Houde ALS, Wilson CC, Neff BD (2015a) Effects of competition with four nonnative salmonid species on Atlantic salmon from three populations. *Transactions of the American Fisheries Society*, **144**, 1081–1090.
- Houde ALS, Wilson CC, Neff BD (2015b) Competitive interactions among multiple non-native salmonids and two populations of Atlantic salmon. *Ecology of Freshwater Fish*, **24**, 44–55.
- Kamada N, Chen GY, Inohara N, Núñez G (2013) Control of pathogens and pathobionts by the gut microbiota. *Nature Immunology*, **14**, 685–690.
- Knowles SR, Nelson EA, Palombo EA (2008) Investigating the role of perceived stress on bacterial flora activity and salivary cortisol secretion: a possible mechanism underlying susceptibility to illness. *Biological Psychology*, **77**, 132–137.
- Kreisinger J, Čížková D, Vohánka J, Piálek J (2014) Gastrointestinal microbiota of wild and inbred individuals of two house mouse subspecies assessed using high-throughput parallel pyrosequencing. *Molecular Ecology*, **23**, 5048–5060.

LeBlanc JG, Milani C, de Giori GS *et al.* (2013) Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current Opinion in Biotechnology*, **24**, 160–168.

Lenz TL, Becker S (2008) Simple approach to reduce PCR artefact formation leads to reliable genotyping of MHC and other highly polymorphic loci--implications for evolutionary analysis. *Gene*, **427**, 117–123.

Ley RE, Peterson DA, Gordon JI (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, **124**, 837–848.

Linnenbrink M, Wang J, Hardouin EA *et al.* (2013) The role of biogeography in shaping diversity of the intestinal microbiota in house mice. *Molecular Ecology*, **22**, 1904–1916.

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, **71**, 8228–8235.

Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL (2005) An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell*, **122**, 107–118.

McKnite AM, Perez-Munoz ME, Lu L *et al.* (2012) Murine gut microbiota is defined by host genetics and modulates variation of metabolic traits. *PLoS ONE*, **7**, e39191.

Merrifield DL, Carnevali O (2014) Probiotic modulation of the gut microbiota of fish. In: *Aquaculture Nutrition* (eds Merrifield D, Ringø E), pp. 185–222. John Wiley & Sons Ltd.

Merrifield DL, Dimitroglou A, Foey A *et al.* (2010) The current status and future focus of probiotic and prebiotic applications for salmonids. *Aquaculture*, **302**, 1–18.

Miller KM, Teffer A, Tucker S *et al.* (2014) Infectious disease, shifting climates, and opportunistic predators: cumulative factors potentially impacting wild salmon declines. *Evolutionary Applications*, **7**, 812–855.

Morantz DL, Sweeney RK, Shirvell CS, Longard DA (1987) Selection of microhabitat in summer by juvenile Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **44**, 120–129.

Neuman C, Hatje E, Zarkasi KZ *et al.* (2016) The effect of diet and environmental temperature on the faecal microbiota of farmed Tasmanian Atlantic salmon (*Salmo salar* L.). *Aquaculture Research*, **47**, 660–672.

O'Hara AM, Shanahan F (2006) The gut flora as a forgotten organ. *EMBO Reports*, **7**, 688–693.

- Oksanen J, Blanchet FG, Kindt R *et al.* (2015) *vegan: Community Ecology Package*. R package version 2.3-1. <http://CRAN.R-project.org/package=vegan>.
- Olszak T, An D, Zeissig S *et al.* (2012) Microbial exposure during early life has persistent effects on natural killer T cell function. *Science*, **336**, 489–493.
- Parks BW, Nam E, Org E *et al.* (2013) Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metabolism*, **17**, 141–152.
- Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics*, **30**, 3123–3124.
- Perdigón G, Fuller R, Raya R (2001) Lactic acid bacteria and their effect on the immune system. *Current Issues in Intestinal Microbiology*, **2**, 27–42.
- Qi Z, Zhang X, Boon N, Bossier P (2009) Probiotics in aquaculture of China — Current state, problems and prospect. *Aquaculture*, **290**, 15–21.
- Ringø E (2008) The ability of carnobacteria isolated from fish intestine to inhibit growth of fish pathogenic bacteria: a screening study. *Aquaculture Research*, **39**, 171–180.
- Ringø E, Løvmo L, Kristiansen M *et al.* (2010) Lactic acid bacteria vs. pathogens in the gastrointestinal tract of fish: a review. *Aquaculture Research*, **41**, 451–467.
- Roesch LFW, Fulthorpe RR, Riva A *et al.* (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal*, **1**, 283–290.
- Scott RJ, Noakes DLG, Beamish FWH, Carl LM (2003) Chinook salmon impede Atlantic salmon conservation in Lake Ontario. *Ecology of Freshwater Fish*, **12**, 66–73.
- Scott RJ, Poos MS, Noakes DLG, Beamish FWH (2005) Effects of exotic salmonids on juvenile Atlantic salmon behaviour. *Ecology of Freshwater Fish*, **14**, 283–288.
- Sharon G, Segal D, Ringo JM *et al.* (2010) Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, **107**, 20051–20056.
- Sjögren K, Engdahl C, Henning P *et al.* (2012) The gut microbiota regulates bone mass in mice. *Journal of Bone and Mineral Research*, **27**, 1357–1367.
- Sogin ML, Morrison HG, Huber JA *et al.* (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proceedings of the National Academy of Sciences*, **103**, 12115–12120.

- Sommer F, Bäckhed F (2013) The gut microbiota--masters of host development and physiology. *Nature Reviews. Microbiology*, **11**, 227–238.
- Stevenson TJ, Buck CL, Duddleston KN (2014) Temporal dynamics of the cecal gut microbiota of juvenile arctic ground squirrels: a strong litter effect across the first active season. *Applied and Environmental Microbiology*, **80**, 4260–4268.
- Stewart TJ, Schaner T (2002) Lake Ontario salmonid introductions 1970 to 1999: stocking, fishery and fish community influences. In: *Lake Ontario Fish Communities and Fisheries: 2001 Annual Report of the Lake Ontario Management Unit*. Queen's Printer for Ontario, pp. 12.1-12.10.
- Sullam KE, Essinger SD, Lozupone CA *et al.* (2012) Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology*, **21**, 3363–3378.
- Van Zwol JA, Neff BD, Wilson CC (2012) The effect of competition among three salmonids on dominance and growth during the juvenile life stage. *Ecology of Freshwater Fish*, **21**, 533–540.
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, **73**, 5261–5267.
- Whitehead A, Triant DA, Champlin D, Nacci D (2010) Comparative transcriptomics implicates mechanisms of evolved pollution tolerance in a killifish population. *Molecular Ecology*, **19**, 5186–5203.
- Wong S, Waldrop T, Summerfelt S *et al.* (2013) Aquacultured rainbow trout (*Oncorhynchus mykiss*) possess a large core intestinal microbiota that is resistant to variation in diet and rearing density. *Applied and Environmental Microbiology*, **79**, 4974–4984.
- Yun JH, Roh SW, Whon TW *et al.* (2014) Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Applied and Environmental Microbiology*, **80**, 5254–5264.
- Zarkasi KZ, Abell GCJ, Taylor RS *et al.* (2014) Pyrosequencing-based characterization of gastrointestinal bacteria of Atlantic salmon (*Salmo salar* L.) within a commercial mariculture system. *Journal of Applied Microbiology*, **117**, 18–27.

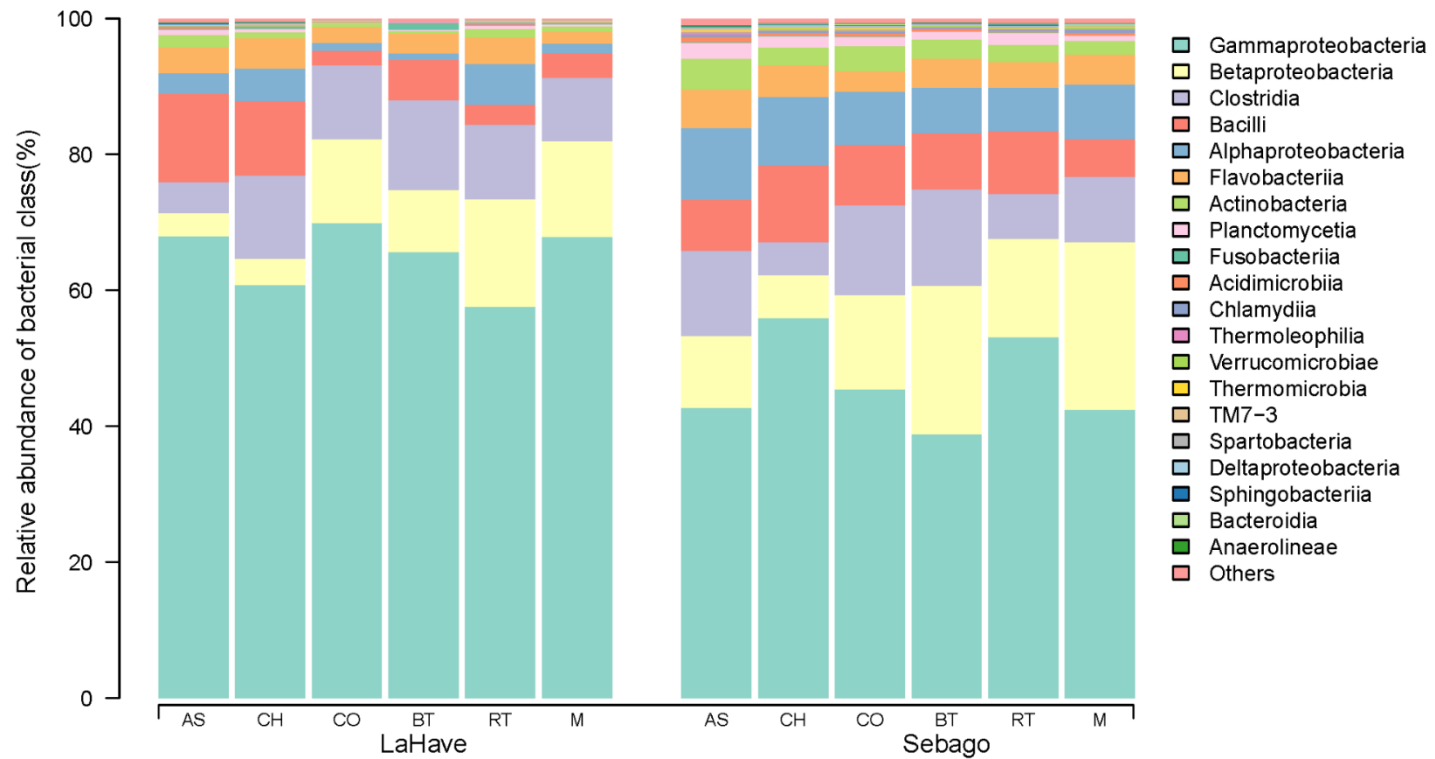
**Table 6.1** Primer sequences used for Atlantic salmon (*Salmo salar*) gut microbiota characterization.

Primer name	Sequences
V5F*	<u>acctgcctgccg</u> ATTAGATACCCNGGTAG
V6R*	<u>acgccaccgagc</u> CGACAGCCATGCANCACCT
UniA†	CCATCTCATCCCTGCGTGTCTCCGACTCAGXXXXXXXXXXGATacctgcctgccg
UniB	CCTCTCTATGGGCAGTCGGTGATacgccaccgagc

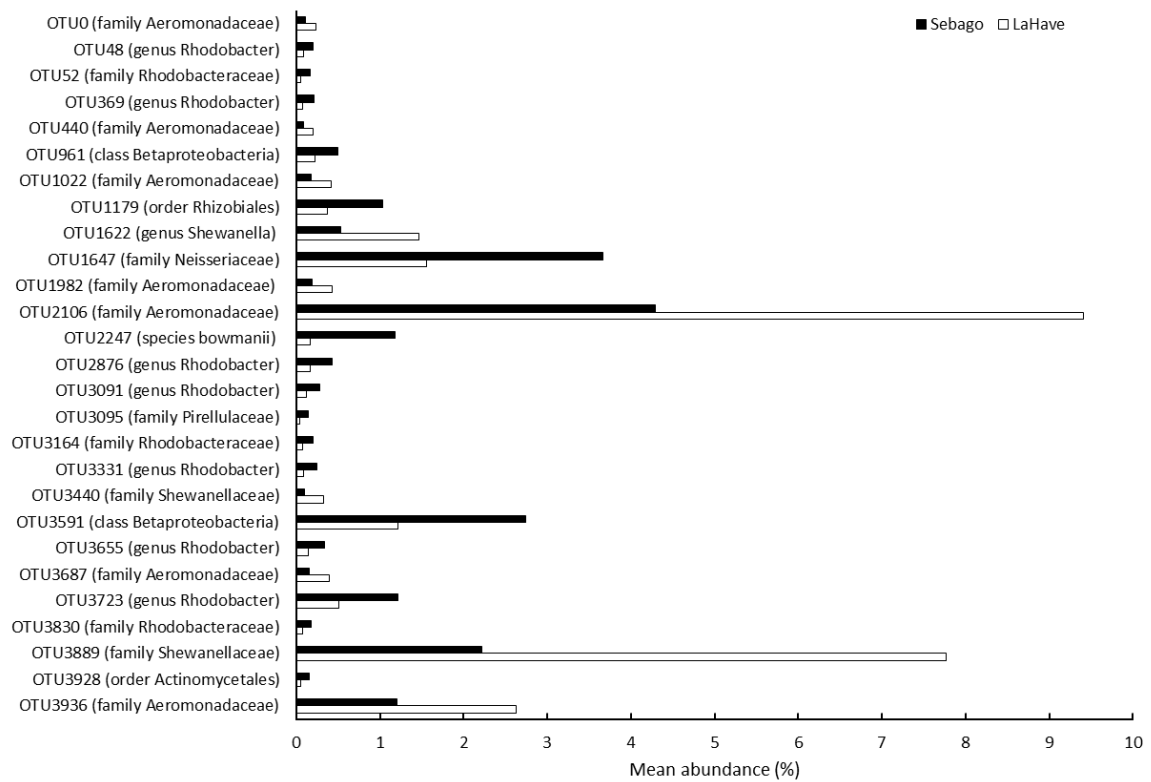
\* The underlined 12 bp sequences in V5F and V6R are tails that bind UniA and UniB in the second PCR respectively.

† The XXXXXXXXXXXX represents different barcode sequences.

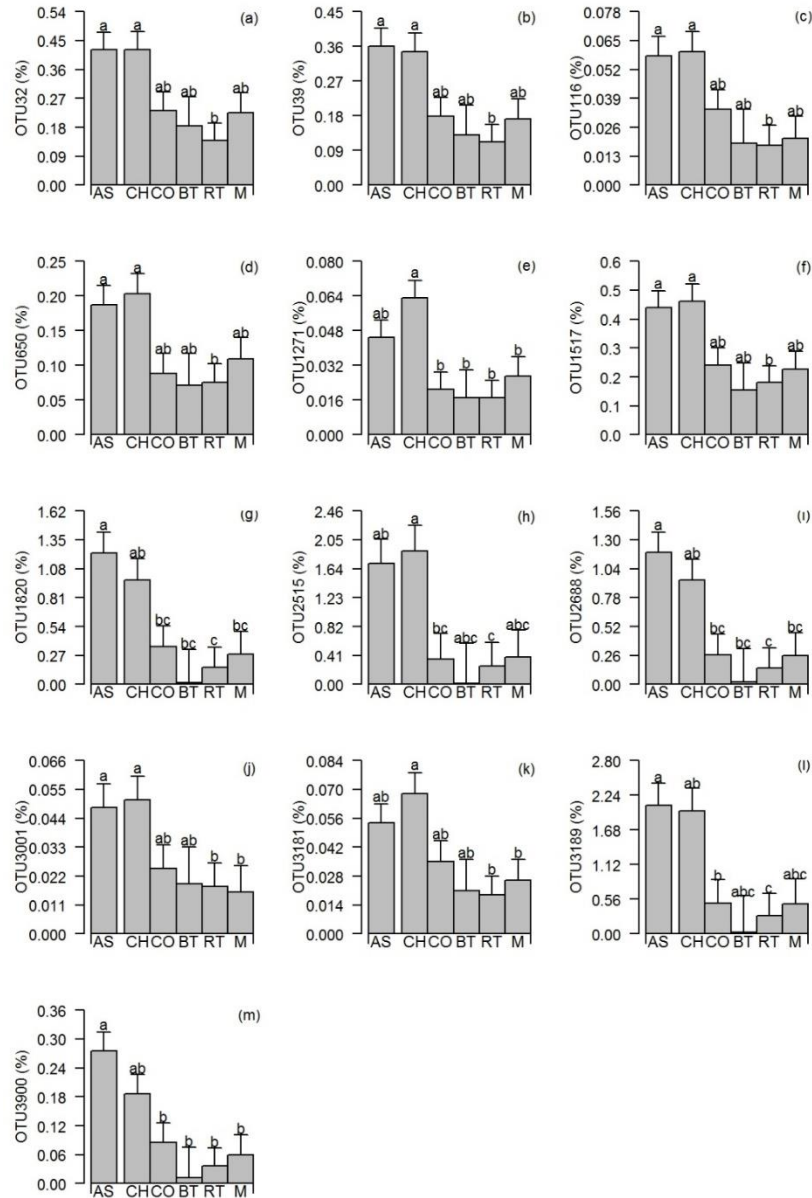




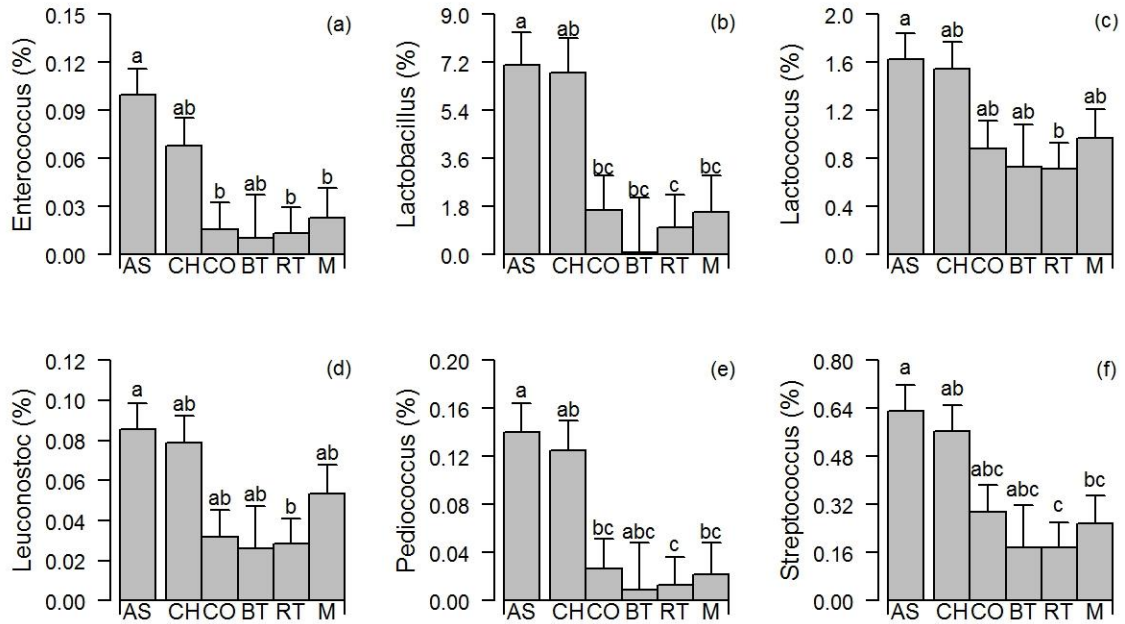
**Figure 6.1** Relative abundance of bacterial classes for juvenile Atlantic salmon (*Salmo salar*) in response to interspecific competition. Displayed are classes with at least 0.1% relative abundance in one treatment. The “others” category includes unclassified sequences at the class level and the sum of all classes that occurred at less than 0.1% relative abundance. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids.



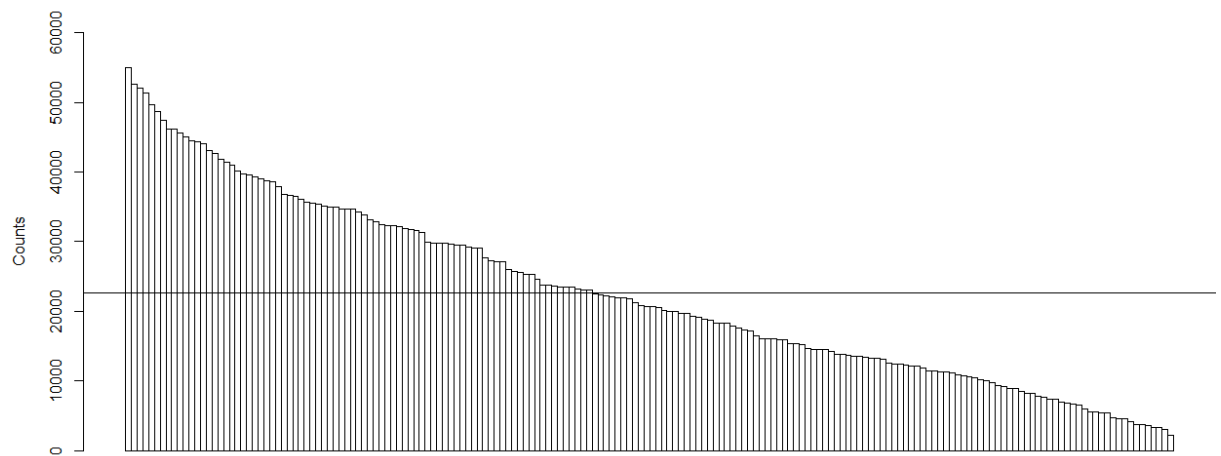
**Figure 6.2** The 27 Operational Taxonomic Units (OTUs) that were significantly different in abundance between the two Atlantic salmon (*Salmo salar*) populations across all treatments. Taxonomic assignment beside each OTU identification number is the lowest taxonomic level obtained.



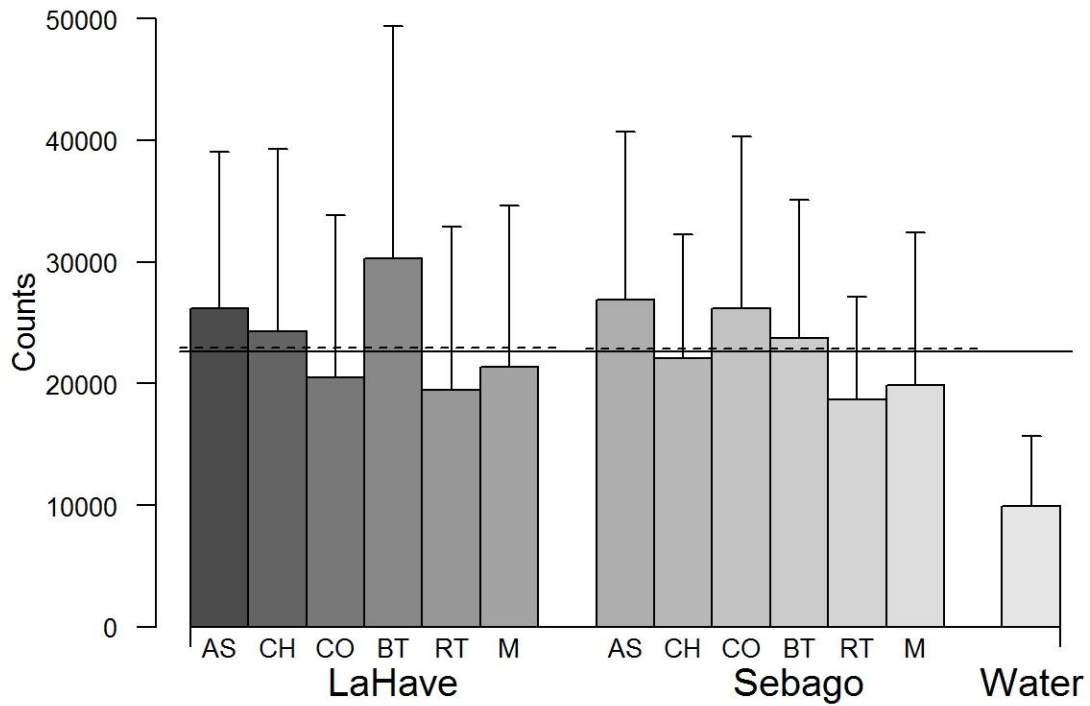
**Figure 6.3** The 13 OTUs showing difference among competition treatments in the Atlantic salmon (*Salmo salar*) LaHave population: OTU32, 116, 650 and 1271 (genus *Chryseobacterium*), OTU39 (genus *Flavobacterium*), OTU1517 (species *succinicans*), OTU1820, 2515 and 3189 (genus *Lactobacillus*), OTU2688 (order *Lactobacillales*), OTU3001 (family *Enterobacteriaceae*), OTU3181 (genus *Wautersiella*), OTU3900 (genus *Streptococcus*). Displayed are means  $\pm$  1 SE for treatments. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. Different letters above the bars indicate significant differences assessed using Tukey's post hoc multiple comparisons ( $P < 0.05$ ).



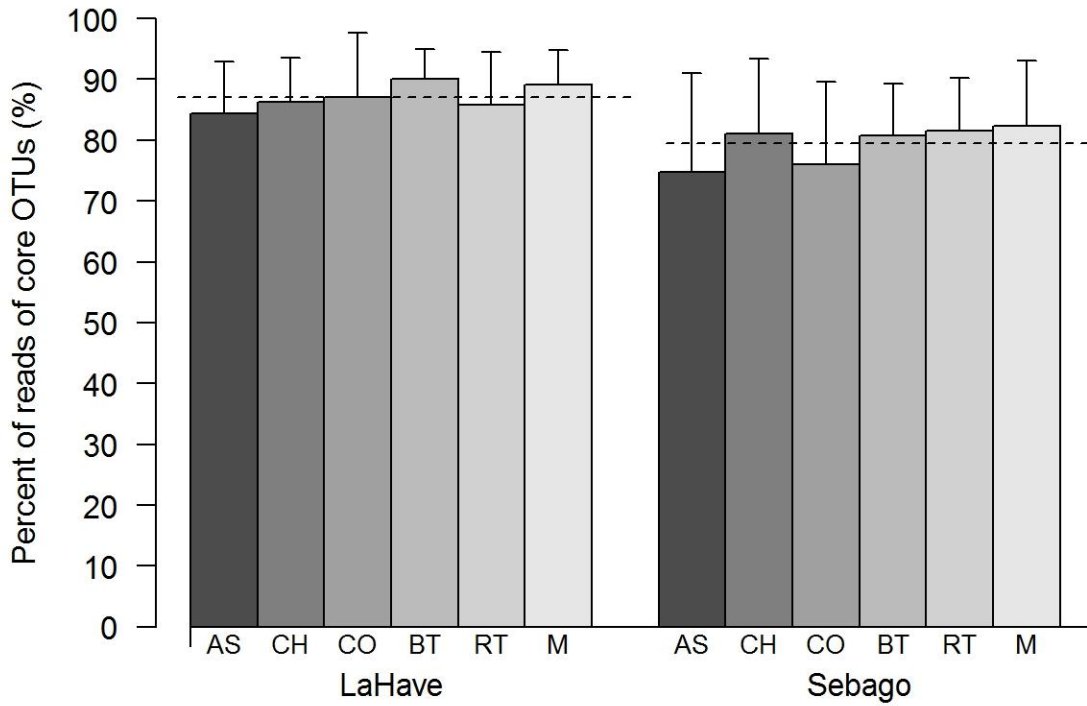
**Figure 6.4** The six lactic acid genera showing differences among treatments in the Atlantic salmon (*Salmo salar*) LaHave population. Displayed are means  $\pm$  1SE for treatments. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids. Different letters above the bars indicate significant differences assessed using Tukey's post hoc multiple comparisons ( $P < 0.05$ ).



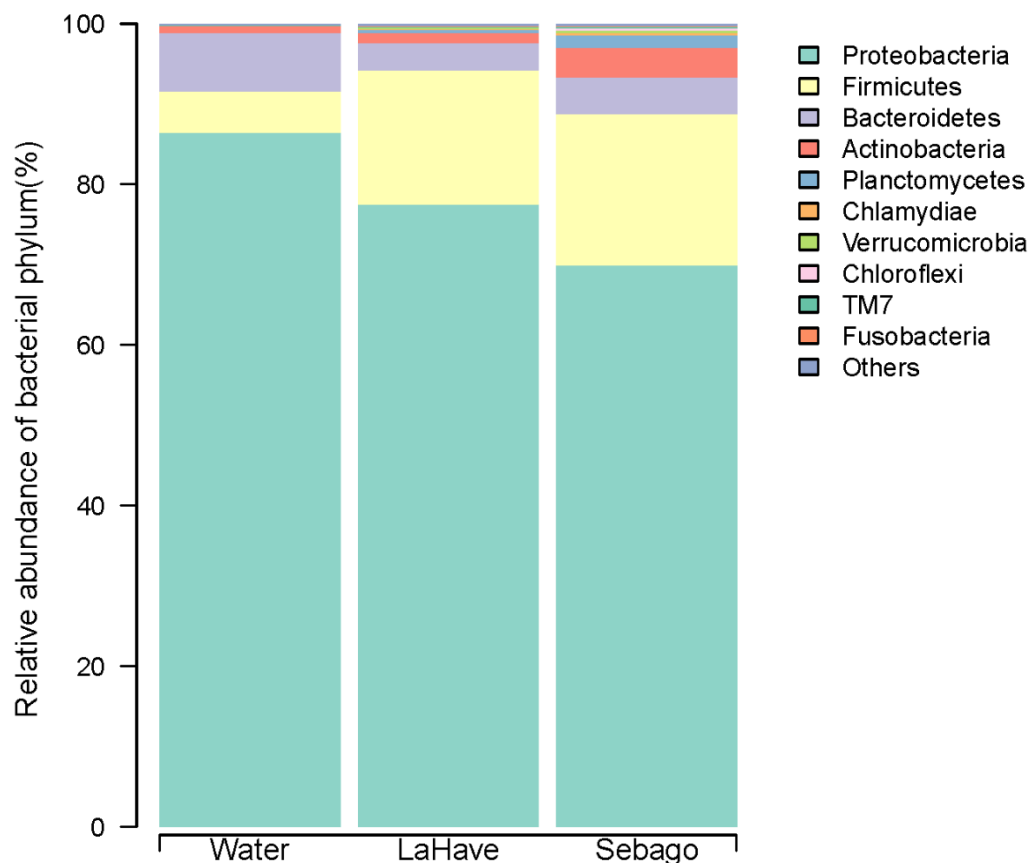
**Supplementary Figure S6.1** Distribution of the number of high quality sequences generated by next generation sequencing of the amplified *16S rRNA* gene for gut microbiota characterization of Atlantic salmon (*Salmo salar*). The solid line represents the average number of sequence reads per sample. Each bar represents one sample.



**Supplementary Figure S6.2** Distribution of the number of high quality sequences generated by next generation sequencing of the amplified *16S rRNA* gene. Displayed are means  $\pm$  1SD for treatments. The solid line represents the average number of sequence reads per sample. The dashed lines represent the average numbers of sequence reads in each of the two Atlantic salmon populations. Treatment symbols: AS indicates Atlantic salmon (*Salmo salar*) reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids.

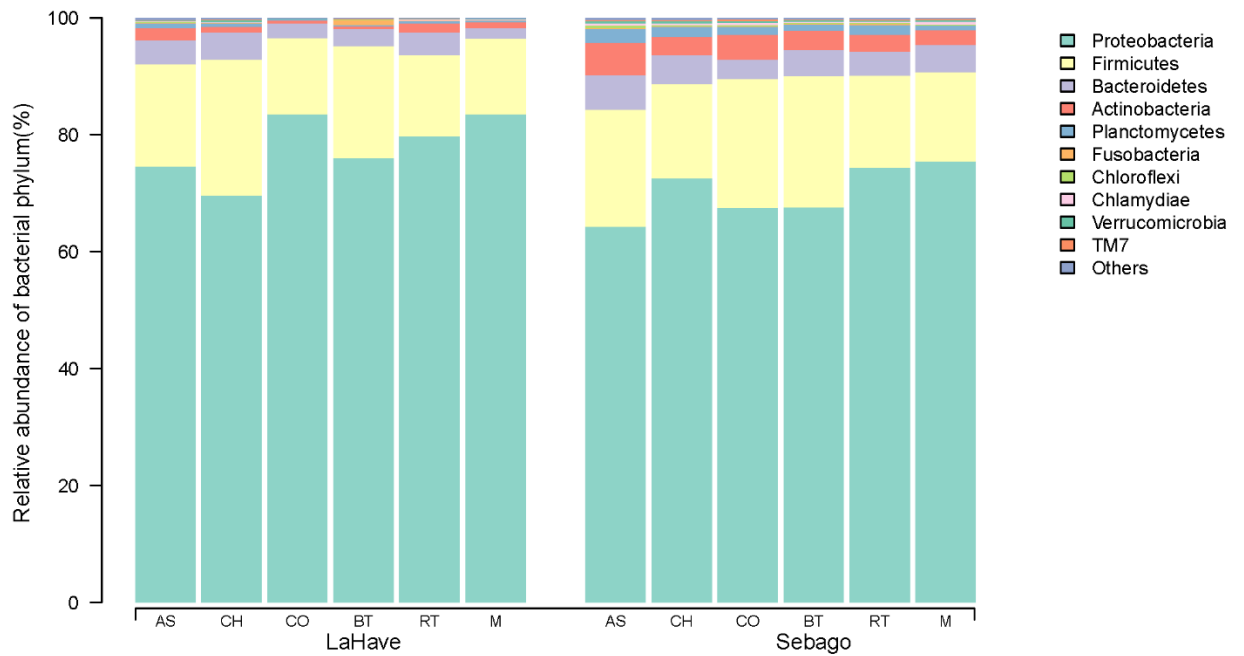


**Supplementary Figure S6.3** The proportion of 16S rRNA sequences accounted by the core OTUs shared by 70% of the gut microbiota samples in Atlantic salmon (*Salmo salar*) for each interspecific competition treatment. Displayed are means  $\pm$  1SD for treatments. The dashed lines represent the average proportion of reads for each of the two populations. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids.

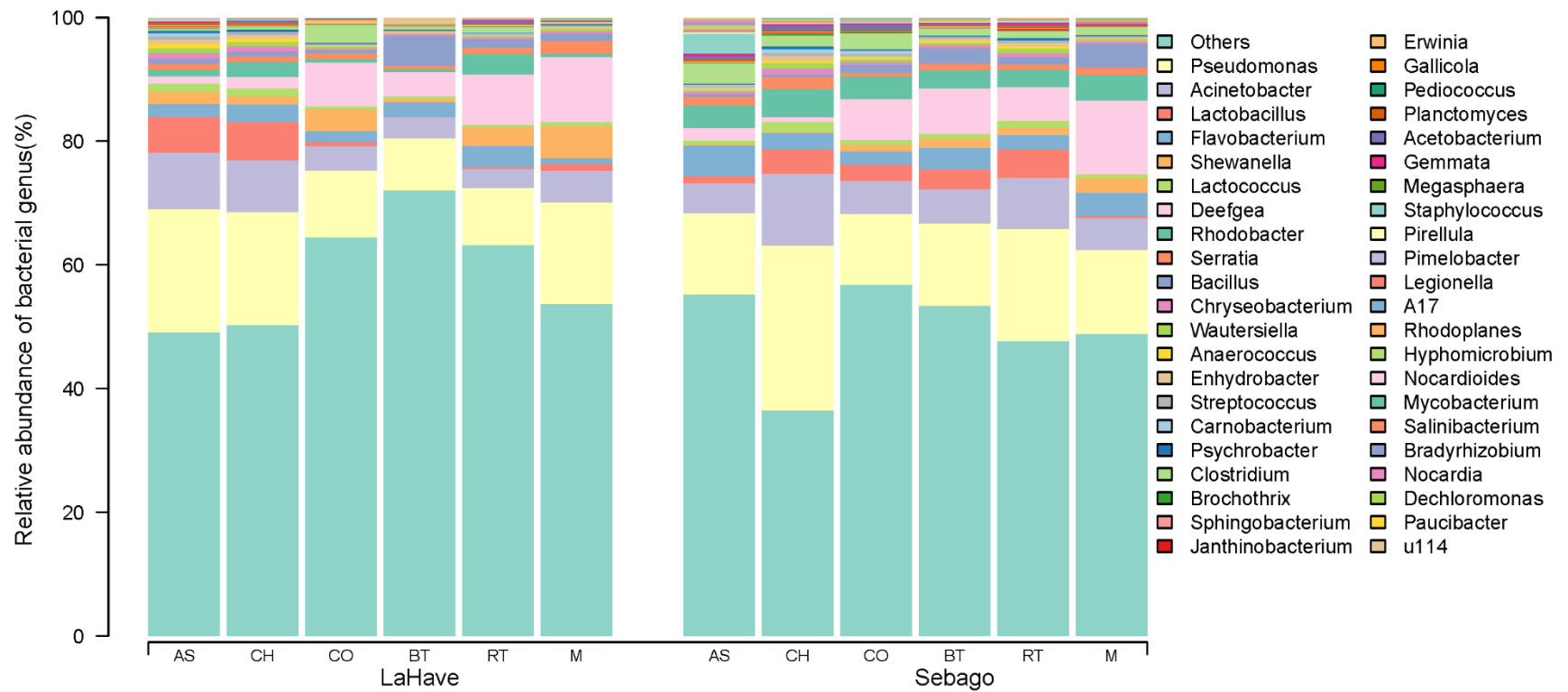


**Supplementary Figure S6.4** Relative abundance of bacterial phyla for the combined water samples (N=4) and the combined samples for all Atlantic salmon (*Salmo salar*) from the LaHave (N=82) and Sebago (N=96) populations. The bars show only phyla with at least 0.1% relative abundance in one of the two Atlantic salmon populations or combined water samples. The “others” category includes unclassified sequences at the phylum level and the sum of all phyla that occurred at less than 0.1% relative abundance.

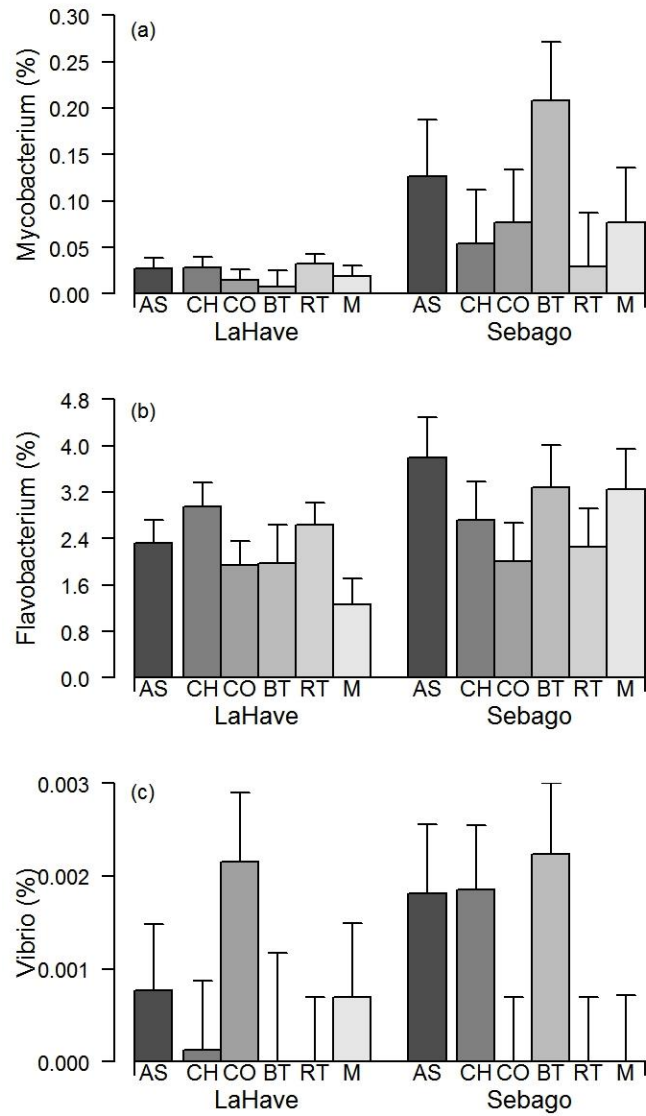




**Supplementary Figure S6.5** Relative abundance of bacterial phyla for juvenile Atlantic salmon (*Salmo salar*) in response to interspecific competition. Displayed are phyla with at least 0.1% relative abundance in one treatment. The “others” category includes unclassified sequences at the phylum level and the sum of all phyla that occurred at less than 0.1% relative abundance. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids.



**Supplementary Figure S6.6** Relative abundance of bacterial genera for juvenile Atlantic salmon (*Salmo salar*) in response to interspecific competition. Displayed are genera with at least 0.1% relative abundance in one treatment. The “others” category includes unclassified sequences at the genus level and the sum of all genera that occurred at less than 0.1% relative abundance. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids.



**Supplementary Figure S6.7** The relative abundance of potential pathogens in the gut microbiota of two Atlantic salmon (*Salmo salar*) populations. There were no significant differences among treatments within each population. The *Aeromonas* genus was undetectable in the data. Displayed are means  $\pm$  1SE for treatments. Treatment symbols: AS indicates Atlantic salmon reared alone; CH, CO, BT, and RT indicate Atlantic salmon reared with one of the four species: Chinook salmon, coho salmon, brown trout, and rainbow trout, respectively; M indicates Atlantic salmon reared with all four non-native salmonids.

## CHAPTER 7

### GENERAL CONCLUSION

#### INTRODUCTION

Successful reintroduction requires first establishment and then long-term persistence (Armstrong & Seddon 2008). Indeed, population differences in establishment performance in novel environments and the corresponding importance of source population selection have long been recognized (Leberg 1993). Thus factors that regulate establishment success in new environments are of importance for more effective management of population reintroductions as well as other conservation-related efforts such as conservation introduction and assisted colonization (Forsman 2014).

Genetic diversity is important to consider when selecting source populations for conservation or commercial purposes (Earnhardt 1999). High genetic variation within a population implies that the population may harbor diverse phenotypes which may be preadapted to new environments or which can be acted on by natural selection (Earnhardt 1999; González-Suárez *et al.* 2015). Genetic diversity is also important for population persistence and long-term success as it helps populations cope with environmental fluctuations (Lande & Shannon 1996). Genetic diversity is commonly measured by neutral genetic markers, but this may not reflect functional genetic variation which is directly relevant to phenotypic variation (Reed & Frankham 2001; Hedrick 2001).

Ecologically significant phenotypes are another important factor to consider in reintroduction as individuals will survive if their phenotypes are close to that supported by the local environment. The phenotypic match may be due to either phenotypic

plasticity or preadapted traits. A recent meta-analysis found that high variation in phenotype can increase establishment success in plants and invertebrates (Forsman 2014). While phenotypic variation may be enhanced by mixing populations, standing intra-population diversity is preferable to avoid outbreeding depression (McClelland & Naish 2006). Although high intra-population variation in phenotype is important, not all phenotypes have fitness consequences: in another meta-analysis, González-Suárez *et al.* (2015) analyzed ten traits and found that only adult body size variation was correlated with establishment success in invasive mammal species. That variation in trait impact on establishment success is especially true in reintroduction efforts when characteristics of the source population must match the environmental conditions of the release site (Houde *et al.* 2015a). For example, Schneider (2011) found that spawning time of Atlantic salmon is a key factor affecting successful reintroduction in the Rhine River. However, key phenotypes which are important for establishment are generally difficult and expensive to identify and characterize for animals, especially fish (Houde *et al.* 2015a).

Gene expression plays an important role in determining phenotypes and in coping with environmental stresses. Gene expression techniques have potential to be used to predict phenotypes (e.g., Miller *et al.* 2011; Tung *et al.* 2012). For instance, Connon *et al.* (2012) found expression of *natural resistance-associated macrophage protein*, *myxovirus resistance*, *chemokine*, and *Cytochrome P450 family 1 subfamily A polypeptide 1* is correlated with different physical conditions in wild rainbow trout (*Oncorhynchus mykiss*). Transcriptomic tools can also be applied to address population differences that may be indicative of preadaptation to environmental factors (Gleason & Burton 2015), and population-specific responses to environmental challenges which may underlie

differences in their tolerance of environmental stress (Whitehead *et al.* 2010). The differentially expressed genes among different phenotypes can in turn be used to predict relevant phenotypes (Miller *et al.* 2011; Tung *et al.* 2012). However, the fact that gene expression is easily influenced by environmental factors may impede the applications of gene expression for prediction.

In this thesis, I explored population differences at the gene expression level (Chapters 3, 4, and 5) and in gut microbiota (Chapter 6). I also explored mechanisms underlying negative effects on Atlantic salmon caused by interspecific competition (Chapters 5 and 6). I found populations exhibited differences in gene expression that are likely due to selection (Chapter 3) and detected significant heritable genetic variation in gene expression variance (Chapter 4). Population-specific responses to interspecific competition occurred at the gene expression level (Chapter 5) and in the composition and diversity of the gut microbiota (Chapter 6). Below I list the major contributions my doctoral research has made to our understanding of population divergence and the impact of gene expression and gut microbiota diversity for conservation efforts.

## CONTRIBUTIONS

### *Population differences in gene expression*

I explored population differences in gene expression in Chapters 3, 4, and 5 to compare transcriptional differences at rest and in response to interspecific competition and to address basic genetic, evolutionary, and ecological theory as it relates to gene expression. In all three chapters, I found gene expression differences between the LaHave and Sebago populations after controlling for rearing environment and age, indicating that

genetic variation at the population level contributes to gene expression profile variation. In Chapter 3, I found population differences in gene expression are likely to be driven by directional selection acting in the local environment and that persisted despite generations in a controlled environment (hatchery). In Chapter 4, I found populations differ in genetic architecture such that genetic variance components of gene transcription showed marked differences between the study populations, indicating expression of the functional genes may show different response to selection after release in the reintroduced sites. In Chapter 5, I compared transcriptomic differences for the two populations at rest and in response to interspecific competition and found that RNA-Seq is an exceptionally sensitive tool to measure stress response and evaluate stress status. Together, these results suggest that gene expression variability is a key factor affecting population differences in fitness and that it has the potential to be used for predicting fitness in new environments.

#### *Population-specific responses to interspecific competition*

I explored the molecular mechanisms underlying the negative effects of interspecific competition on Atlantic salmon in Chapters 5 and 6, and found that fish from the two populations showed surprisingly different responses to interspecific competition. In Chapter 5, I found the magnitude of the transcriptional responses were smaller for both populations of Atlantic salmon in competition with Chinook salmon than in competition with rainbow trout and brown trout. I also found that the LaHave population showed substantial transcriptional responses to competition with rainbow trout and brown trout, whereas only competition with brown trout caused substantial transcriptional responses for the Sebago population. In Chapter 6, I found that

interspecific competition reduced the relative abundance of 13 OTUs and lactic acid bacteria (beneficial bacteria) in the gut microbiotic community of the LaHave population but not the Sebago population. My work is the first to show consistent microbial and transcriptional population differences in response to interspecific competition, and it highlights that those effects can have important implications in conservation and reintroduction biology.

There were similarities and differences for my results from Chapters 5 and 6 in comparison to previously published results on the effects of interspecific competition on fitness-related traits for the Sebago and LaHave populations. In both chapters, I found little response to rearing with Chinook salmon; this was expected, as previous studies had found that Chinook salmon had no negative effects on the fitness-related traits of Atlantic salmon because of similar levels of aggression between species (Houde *et al.* 2015b; c). However, in Chapter 5, my transcriptional results indicated that rainbow trout may be the most stressful competitor for the LaHave Atlantic salmon and brown trout may be the most stressful competitor for the Sebago Atlantic salmon; this is different from previously published work based on fitness-related traits which concluded that the three species are ranked brown trout, rainbow trout, and Atlantic salmon in order of dominance (Van Zwol *et al.* 2012). In Chapter 6, I found that Atlantic salmon in competition with coho salmon showed a pattern of gut microbiotic composition similar to that of Atlantic salmon reared with more aggressive species (i.e. brown trout and rainbow trout). This was surprising because Atlantic salmon is thought to be just as aggressive as coho salmon (Houde *et al.* 2015b; c). Although my results generally agreed with previous results based on fitness-related traits, the important discrepancies and the higher divergence in



response between the two populations in my research indicate that molecular biological techniques are more sensitive than fitness-related traits to environmental stress.

#### *Source population selection for reintroduction in Lake Ontario*

Based on the results from Chapters 5 and 6, I concluded that the Sebago population is more appropriate for reintroduction into Lake Ontario than the LaHave population. In Chapters 5 and 6, I found that interspecific competition led to more and larger changes at both the gene expression level and in the gut microbiota composition in the LaHave population than in the Sebago population, indicating that the Sebago population has higher competitive ability or higher tolerance to interspecific competition. As interspecific competition with non-native salmonid is thought to be an impediment to the successful reintroduction of Atlantic salmon in Lake Ontario (Scott *et al.* 2003, 2005; Van Zwol *et al.* 2012; Houde *et al.* 2015b; c, 2016), the Sebago population should be a more suitable population for reintroduction in Lake Ontario.

### FUTURE DIRECTIONS

My doctoral work provides fundamental contributions to reintroduction biology and more broadly conservation biology. My work impacts both applied and basic science and also contributed substantially to the technical tools available for ecologists, conservation biologists, and evolutionary biologists. Furthermore, my work has also highlighted a number of exciting and important future directions.

First, interspecific competition with non-native salmonids is thought to be a major barrier to the successful reintroduction of Atlantic salmon into Lake Ontario (Scott *et al.*

2003; Van Zwol *et al.* 2012; Houde *et al.* 2015b; c, 2016), and interspecific competition in artificial stream tanks is generally more intensive than in natural streams (Korsu *et al.* 2010). I studied population differences in response to interspecific competition (Chapters 5 and 6) using juvenile Atlantic salmon in artificial stream tanks. In the future, it is critical to expand on my work by releasing Atlantic salmon into Lake Ontario tributaries where non-native salmonids are both present and absent, and re-capture the fish to quantify their response at the gene expression level and in their gut microbiota. It would also be valuable to compare more populations in their response to interspecific completion at a greater range of development stages.

Second, gut microbiota play an important role in the health of their host through complex interactions between the microbiota and the host (Carabotti *et al.* 2015). In fact, the gut microbiota has been proposed as a potential biomarker for type 2 diabetes and cardio-metabolic diseases in human clinical studies (Vinjé *et al.* 2014; Yassour *et al.* 2016). In the future, gut microbiotic composition should be investigated in response to a variety of environmental stressors. If populations with higher gut bacteria diversity are consistently more tolerant to stressors, characterizing gut microbiota could aid in source population selection. However, how gut microbiota respond to different kinds of stressors and the function of gut microbiota composition and diversity should be further investigated before its application in conservation management.

Third, in addition to biotic stressors, abiotic stressors, such as temperature and pollutant exposure, have contributed to the decline and extirpation of fish species worldwide (Snucins *et al.* 1995; Parrish *et al.* 1998; Xenopoulos *et al.* 2005; Wenger *et al.* 2011), thus variation in tolerance to abiotic stressors can affect the establishment of

organisms in new environments. This is especially true for fish populations used for reintroduction, as captive fish populations are reared in conditions with consistent temperature and water quality.

Lastly, my studies focuses on population establishment of reintroduction. In the future, it is important to measure genetic diversity of Atlantic salmon source populations using functional loci as I proposed in Chapter 2 to rank adaptive potential of source populations. The combination of monitoring performance of Atlantic salmon populations after release into Lake Ontario tributaries and the application of genomic tools to estimate population differences in persistence will not only provide more information for source population selection for reintroduction, but also lead to better understanding of fitness-related genetic variation in salmonids.

## REFERENCES

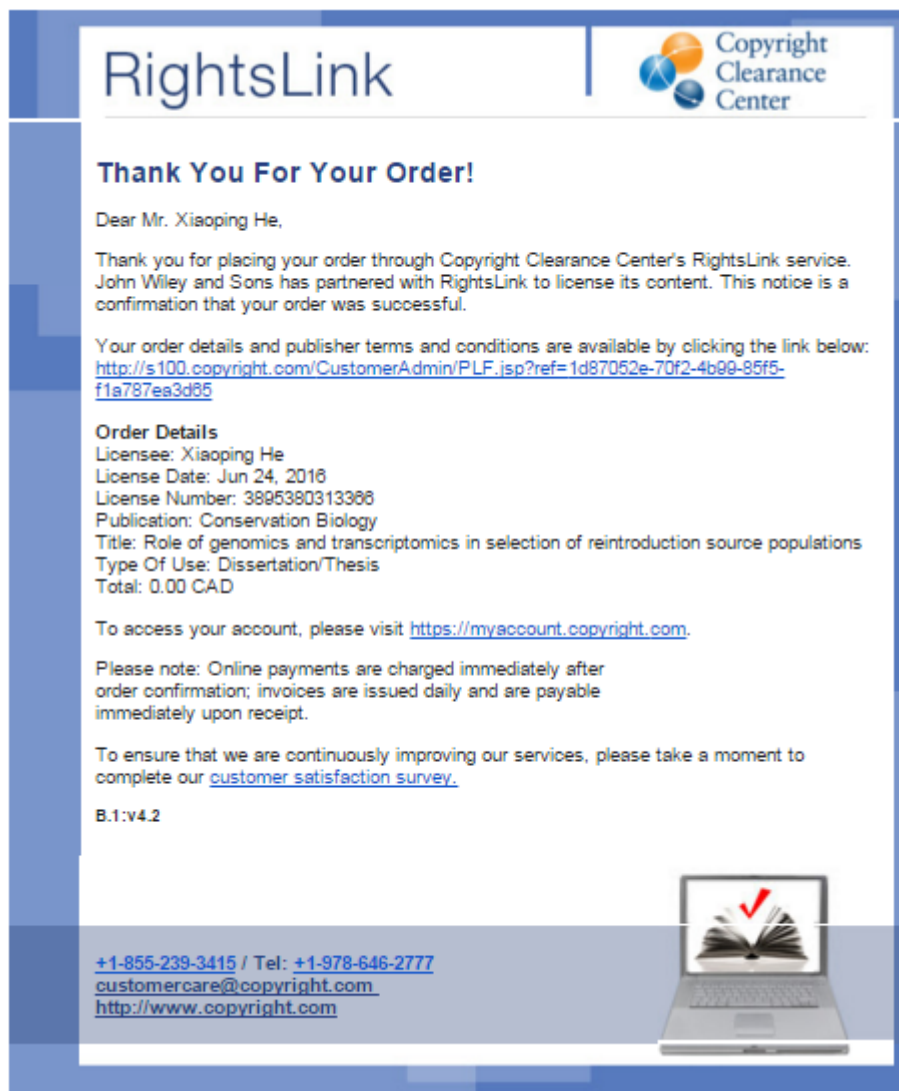
- Armstrong DP, Seddon PJ (2008) Directions in reintroduction biology. *Trends in Ecology & Evolution*, **23**, 20–25.
- Carabotti M, Scirocco A, Maselli MA, Severi C (2015) The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Annals of Gastroenterology*, **28**, 203–209.
- Connon RE, D’Abronzio LS, Hostetter NJ *et al.* (2012) Transcription profiling in environmental diagnostics: health assessments in Columbia River basin steelhead (*Oncorhynchus mykiss*). *Environmental Science & Technology*, **46**, 6081–6087.
- Earnhardt JM (1999) Reintroduction programmes: genetic trade-offs for populations. *Animal Conservation*, **2**, 279–286.
- Forsman A (2014) Effects of genotypic and phenotypic variation on establishment are important for conservation, invasion, and infection biology. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 302–307.
- Gleason LU, Burton RS (2015) RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail *Chlorostoma funebris*. *Molecular Ecology*, **24**, 610–627.
- González-Suárez M, Bacher S, Jeschke JM (2015) Intraspecific trait variation is correlated with establishment success of alien mammals. *The American Naturalist*, **185**, 737–746.
- Hedrick PW (2001) Conservation genetics: Where are we now? *Trends in Ecology & Evolution*, **16**, 629–636.
- Houde ALS, Garner SR, Neff BD (2015a) Restoring species through reintroductions: strategies for source population selection. *Restoration Ecology*, **23**, 746–753.
- Houde ALS, Smith AD, Wilson CC, Peres-Neto PR, Neff BD (2016) Competitive effects between rainbow trout and Atlantic salmon in natural and artificial streams. *Ecology of Freshwater Fish*, **25**, 248–260.
- Houde ALS, Wilson CC, Neff BD (2015b) Competitive interactions among multiple non-native salmonids and two populations of Atlantic salmon. *Ecology of Freshwater Fish*, **24**, 44–55.

- Houde ALS, Wilson CC, Neff BD (2015c) Effects of competition with four nonnative salmonid species on Atlantic salmon from three populations. *Transactions of the American Fisheries Society*, **144**, 1081–1090.
- Korsu K, Huusko A, Muotka T (2010) Impacts of invasive stream salmonids on native fish: using meta-analysis to summarize four decades of research. *Boreal Environment Research*, **15**, 491–500.
- Lande R, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, **50**, 434–437.
- Leberg PL (1993) Strategies for population reintroduction: effects of genetic variability on population growth and size. *Conservation Biology*, **7**, 194–199.
- McClelland EK, Naish KA (2006) What is the fitness outcome of crossing unrelated fish populations? A meta-analysis and an evaluation of future research directions. *Conservation Genetics*, **8**, 397–416.
- Miller KM, Li S, Kaukinen KH *et al.* (2011) Genomic signatures predict migration and spawning failure in wild Canadian salmon. *Science*, **331**, 214–217.
- Parrish DL, Behnke RJ, Gephard SR, McCormick SD, Reeves GH (1998) Why aren't there more Atlantic salmon (*Salmo salar*)? *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 281–287.
- Reed DH, Frankham R (2001) How closely correlated are molecular and quantitative measures of genetic variation? A meta-analysis. *Evolution*, **55**, 1095–1103.
- Schneider J (2011) Review of reintroduction of Atlantic salmon (*Salmo salar*) in tributaries of the Rhine River in the German Federal States of Rhineland-Palatinate and Hesse. *Journal of Applied Ichthyology*, **27**, 24–32.
- Scott RJ, Noakes DLG, Beamish FWH, Carl LM (2003) Chinook salmon impede Atlantic salmon conservation in Lake Ontario. *Ecology of Freshwater Fish*, **12**, 66–73.
- Scott RJ, Poos MS, Noakes DLG, Beamish FWH (2005) Effects of exotic salmonids on juvenile Atlantic salmon behaviour. *Ecology of Freshwater Fish*, **14**, 283–288.
- Snucins EJ, Gunn JM, Keller W (1995) Restoration of the Aurora trout to its acid-damaged native habitat. *Conservation Biology*, **9**, 1307–1311.
- Tung J, Barreiro LB, Johnson ZP *et al.* (2012) Social environment is associated with gene regulatory variation in the rhesus macaque immune system. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 6490–6495.

- Van Zwol JA, Neff BD, Wilson CC (2012) The effect of competition among three salmonids on dominance and growth during the juvenile life stage. *Ecology of Freshwater Fish*, **21**, 533–540.
- Vinje S, Stroes E, Nieuwdorp M, Hazen SL (2014) The gut microbiome as novel cardio-metabolic target: the time has come! *European Heart Journal*, **35**, 883–887.
- Wenger SJ, Isaak DJ, Luce CH *et al.* (2011) Flow regime, temperature, and biotic interactions drive differential declines of trout species under climate change. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 14175–14180.
- Whitehead A, Triant DA, Champlin D, Nacci D (2010) Comparative transcriptomics implicates mechanisms of evolved pollution tolerance in a killifish population. *Molecular Ecology*, **19**, 5186–5203.
- Xenopoulos MA, Lodge DM, Alcamo J *et al.* (2005) Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology*, **11**, 1557–1564.
- Yassour M, Lim MY, Yun HS *et al.* (2016) Sub-clinical detection of gut microbial biomarkers of obesity and type 2 diabetes. *Genome Medicine*, **8**, 17.


## APPENDIX A: REPRINT PERMISSIONS

### Reprint permission for Chapter 2



## Reprint permission for Chapter 3

**RightsLink**

 Copyright  
Clearance  
Center

**Thank You For Your Order!**

Dear Mr. Xiaoping He,

Thank you for placing your order through Copyright Clearance Center's RightsLink service. Springer has partnered with RightsLink to license its content. This notice is a confirmation that your order was successful.

Your order details and publisher terms and conditions are available by clicking the link below:  
<http://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=d1ef7154-807d-4842-8b05-61a2088ae9ae>

**Order Details**  
Licensee: Xiaoping He  
License Date: Jun 23, 2016  
License Number: 3894910135877  
Publication: Conservation Genetics  
Title: Transcriptional profiling of two Atlantic salmon strains: implications for reintroduction into Lake Ontario  
Type Of Use: Thesis/Dissertation  
Total: 0.00 CAD


To access your account, please visit <https://myaccount.copyright.com>.

Please note: Online payments are charged immediately after order confirmation; invoices are issued daily and are payable immediately upon receipt.

To ensure that we are continuously improving our services, please take a moment to complete our [customer satisfaction survey](#).

B.1:v4.2

[+1-855-239-3415](tel:+18552393415) / Tel: [+1-978-646-2777](tel:+19786462777)  
[customercare@copyright.com](mailto:customercare@copyright.com)  
<http://www.copyright.com>





## VITA AUCTORIS

NAME:	Xiaoping He
PLACE OF BIRTH:	Honghu, Hubei, China
YEAR OF BIRTH:	1984
EDUCATION:	Huazhong Agricultural University, B.Sc., Wuhan, Hubei, China, 2005  Huazhong Agricultural University, M.Sc., Wuhan, Hubei, China, 2008