From Genes to Communities: Effects of Habitat Change over Space and Time on Fish Diversity

by

Pasan Samarasin-Dissanayake

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Ecology and Evolutionary Biology
University of Toronto

© Copyright by Pasan Samarasin-Dissanayake 2015
Abstract

Rapid habitat changes, caused by human activities within the last century, have resulted in biodiversity loss at all levels of biological organization. In light of these changes in the recent past, this thesis explores some of the effects of such changes on fish and wildlife populations, species, and communities using theoretical and empirical approaches. Using simulated genetic data, I first investigated the effects of recent population connectivity changes on the reliability of genetic inferences about connectivity. I found that, when connectivity has declined in the recent past, commonly-used genetic methods for estimating connectivity tend to overestimate current connectivity and underestimate historical connectivity. This could lead to incorrect inferences about gene flow and have negative consequences for conservation of populations and species at risk.

Next, I conducted two empirical studies focusing on Canadian fishes that are threatened by habitat changes, making them excellent systems to investigate the effects of habitat change on ecology and evolution. For Sockeye salmon populations in the Fraser River, I found that hydroelectric dams have fragmented habitats, changed connectivity among populations, and had
significant effects on the ecology and evolution of some populations. Based on molecular data, I found evidence indicating very early differentiation between anadromous and resident forms of Sockeye salmon within one reservoir, where a dam has prevented the historically anadromous salmon from migrating to the ocean.

In a second case study, on freshwater fish communities in northern Canadian lakes that were thought to be depauperate in biodiversity, I found higher species diversity and fish biomass than expected based on species-energy theory. My analyses indicate that fish diversity and biomass in northern lakes are not substantially lower than southern Canadian lakes. Thus, northern lakes could be important reserves of coldwater fish in light of climate change.

In summary, I examined some effects of habitat change on populations, species and communities, and my thesis highlights: (i) areas for the improvement of methodologies and inferences used by conservation biologists; and (ii) specific northern communities of Canadian fishes that warrant attention to preserve Canadian fish diversity.
Acknowledgments

First and foremost, I would like to thank my advisors, Dr. Helen Rodd and Dr. Brian Shuter, for helping me get to this point. Without their assistance and taking me in as a student at a difficult time for me, I would not have reached the finish line. Therefore, I am eternally grateful for both of them for their support and guidance in this journey. I greatly admire their openness and willingness to help at all times, which is what I will take with me the most from them. I was truly blessed to have such kind advisors as Helen & Brian. I am also grateful for my committee members, Dr. Stephen Wright and Dr. Ken Minns for their guidance with research and encouragement. Stephen’s help with two chapters related to conservation genetics was invaluable. Ken was always available, fast with feedback and providing guidance on research and careers. I would like to thank Dr. Paul Bentzen and Dr. Nicholas Mandrak for providing their time to serve in the examination committee. I am thankful to Dr. Allan Baker, my MSc. advisor, who recently passed away, provided support at a difficult time for me. It is with sadness that I am recalling my meetings with him, his kind words and willingness to help at a time of need. Many other EEB faculty members have helped me during my PhD tenure especially Dr. Don Jackson, Dr. Marie-Josee Fortin, Dr. Nick Collins, Dr. Ben Gilbert, Dr. Helene Wagner, Dr. Becky Raboy, and Dr. Mart Gross. I am also thankful to Kitty Lam, Jenn English, and EEB administrative staff for their help with administrative and financial matters for all these years. Thank you to Jill Wheeler, Dominic Fench, Mary Ann Honig, and Carol Finley for helping me to teach BIO150/120/220 labs for many years.

Many graduate students (past and present) at the department have given me moral support, research assistance and occasional fun. Robert Williamson and Emily Josephs helped me with Python programming. I am thankful to Aaron Hall for all sorts of discussions (many over lunch),
Alex DeSerrano for much needed fun and silliness, Liset Cruz-Font for giving a helping hand, Crystal Vincent for providing humour, and many office and lab mates including Anna Li, Mitch Daniels, Lucia Kwan, Jordan Pleet, David Punzalan, Ilona Naujokaitis-Lewis, and Brie Edwards for their various contributions.

I am extremely thankful to my parents and my brother for their support during this time. It would not have been possible to reach the finish line without their encouragement and all modes of support – moral and otherwise. Thanks to the numerous times home-cooked food that got delivered even without asking and rides home by my brother. I am grateful for their love and also not annoying me by constantly asking about when I will finish. I also thank my sister for keeping me in her thoughts and wishing the best for me from afar. Finally, I am extremely grateful to my fiancée Alivia Dey for love, support, patience, advice and encouragement throughout. I am thankful for her cooking and baking skills that kept me very well fed without the need for junk food. I am thankful to her for also pushing me to have some fun through socializing and travelling. I look forward to more travelling and fun with you.
# Table of Contents

Acknowledgments........................................................................................................................... ii

Table of Contents........................................................................................................................... vi

List of Tables ................................................................................................................................. xi

List of Figures ............................................................................................................................... xii

List of Appendices ........................................................................................................................ xv

Chapter 1 ......................................................................................................................................... 1

Introduction..................................................................................................................................... 1

1.1 General background .............................................................................................................. 1

1.2 Scale considerations in conservation................................................................................ 1

1.3 Habitat changes in space and time ................................................................................... 2

1.4 Status of Canadian fishes ................................................................................................. 4

1.5 Thesis objectives and motivation ..................................................................................... 5

1.6 Literature cited ..................................................................................................................... 8

Chapter 2 ....................................................................................................................................... 11

Genetic connectivity in a changing world: How accurately can we estimate current genetic connectivity when population connectivity has changed in recent times? ............................................................... 11

2.1 Abstract .............................................................................................................................. 11

2.2 Introduction ......................................................................................................................... 12
Chapter 3 ....................................................................................................................................... 46

After 100 years: Hydroelectric dam-induced life-history divergence and population genetic changes in Sockeye salmon ............................................................................................................................................ 46

3.1 Abstract ................................................................................................................................ 46

3.2 Introduction ........................................................................................................................ 47

3.3 Materials and Methods .................................................................................................... 50

3.3.1 DNA sampling ........................................................................................................ 50

3.3.2 Measures of genetic diversity and differentiation ................................................... 51

3.3.3 Genetic bottlenecks ................................................................................................. 52

3.3.4 Population genetic structure .................................................................................... 53

3.3.5 Genetic connectivity ............................................................................................... 54

3.3.6 Origins of sea-run and resident types...................................................................... 56

3.4 Results ............................................................................................................................ 57

3.4.1 Genetic diversity and differentiation ...................................................................... 57

3.4.2 Genetic bottlenecks ................................................................................................. 57

3.4.3 Population genetic structure: all sites ..................................................................... 58

3.4.4 Population genetic structure: Alouette and Coquitlam individuals ....................... 59

3.4.5 Population splitting time estimate ........................................................................... 59

3.4.6 Long-term historical connectivity vs. contemporary connectivity ......................... 60

3.5 Discussion ...................................................................................................................... 61
3.5.1 Early-stage ecological divergence in sympatry? ..................................................... 61
3.5.2 Past and present genetic connectivity among all eleven populations ................. 64
3.5.3 Conservation implications ...................................................................................... 66
3.5.4 Conclusions ............................................................................................................. 66

3.6 Acknowledgements ........................................................................................................ 67
3.7 Literature cited ............................................................................................................... 68
3.8 Tables ............................................................................................................................. 74
3.9 Figures ............................................................................................................................ 78

Chapter 4 ....................................................................................................................................... 85

Fish diversity and biomass in northern Canadian lakes: northern lakes are more diverse and have greater biomass than expected based on species-energy theory ................................................... 85

4.1 Abstract .......................................................................................................................... 85
4.2 Introduction .................................................................................................................... 86
4.3 Materials and methods ................................................................................................... 89
  4.3.1 Study area and data ................................................................................................. 89
  4.3.2 Analyses .................................................................................................................. 90
4.4 Results ............................................................................................................................ 94
  4.4.1 Species – area relationship ...................................................................................... 94
  4.4.2 Fish abundance and total biomass ........................................................................... 95
  4.4.3 Dominant species .................................................................................................... 96
# Table of Contents

4.5  Discussion ...................................................................................................................... 97

4.5.1  Biomass ................................................................................................................... 99

4.5.2  Dominant species .................................................................................................. 100

4.5.3  Future of northern fishes ....................................................................................... 102

4.6  Acknowledgements ...................................................................................................... 103

4.7  Literature cited ............................................................................................................. 103

4.8  Table ............................................................................................................................. 112

4.9  Figures .......................................................................................................................... 114

4.10  Appendix ...................................................................................................................... 122

4.11  Copyright Acknowledgement ...................................................................................... 127

Chapter 5 ..................................................................................................................................... 128

Conclusions ................................................................................................................................. 128

5.1  Future directions ........................................................................................................... 131

5.1.1  Understanding effects of recent connectivity changes on genetic inferences ...... 131

5.1.2  Evolution of life-history types in Sockeye salmon ............................................... 132

5.1.3  Quantifying aquatic resources in the northern Canada ......................................... 132

5.2  Final remarks ................................................................................................................ 133

5.3  Literature cited ............................................................................................................. 134
List of Tables

Table 2.1: The ratio of mean square error when connectivity change was 5 generation ago (i.e., recent) versus when connectivity change was 500 generations ago (i.e., historic). .......................... 33

Table 3.1: Measures of genetic diversity for 11 putative Sockeye salmon populations from Lower Fraser River. Abbreviations are as follows: n = sample size, k = mean number of alleles, $H_o$ = observed gene diversity within individuals, $H_e$ = expected gene diversity among individuals, $N_e$ = effective population size estimate (± standard error of the mean), $P$ Bottleneck = Probability of bottleneck using one-tailed Wilcoxon sign rank test implemented in BOTTLENECK v.1.2.02, assuming 90% stepwise mutations. Bold font indicates a statistically significant probability of a bottleneck. ..................................................................................................................................... 74

Table 3.2: Pairwise Fst estimates between 11 putative Sockeye salmon populations estimated using 14 microsatellite loci. Bold values indicate significance (P< 0.00091) after Bonferroni correction. ..................................................................................................................................... 75

Table 3.3: Estimates of gene flow ($N_e m$) among lower Fraser River Sockeye salmon populations, based on a) Migrate-n analysis, and b) BayesAss analysis. .............................................................................................................. 76

Table 4.1: Summary of main analyses and information on data used for the analysis ............... 112
List of Figures

Figure 2.1: A schematic showing an example of connectivity among four populations of a hypothetical species pre and post road construction. a) Four populations exchange migrants with no barrier to restrict migration among patches (prior to road construction). b) Roads were constructed 50 generations ago and effective migration was reduced to zero from that time until present day. ................................................................. 34

Figure 2.2: Plot of $F_{st}$ over 5000 generations for a data set with population size of 200 ($N$=200) and a migration rate change from $Nm=1$ to $Nm=0$, 50 generations before the end ($Nm_{before} = 1$, $Nm_{after} = 0$). Note the increase in $F_{st}$ in the last 50 generations as a result of the reduction in migration. ............................................................................................................. 35

Figure 2.3: The mean square error (panel a) and bias (panel b) of migration rate estimates using BayesAss (dark bar) and Migrate-n (grey bar) when gene flow changed from $Nm=1$ to $Nm=0$ at different points in time (x-axis) with different population sizes ($N=100$, 200, 1000, 2000). Error bars represent the standard error of the mean. ................................................................................................................. 36

Figure 2.4: Relationship between relative mean square error (RMSE) and magnitude of gene flow change ($\Delta Nm$) for population size of 200 ($N=200$). Error bars represent the standard error of the mean. The figure shows increasing accuracy (i.e., decreasing RMSE) as the magnitude of the gene flow change decrease.......................................................................................................................... 37

Figure 2.5: Effect of time since the change in connectivity on relative mean square error when time is scaled to the population size used in the simulation. A gene flow change from $Nm =1$ to $Nm= 0.25$ was used to generate the results for this figure.......................................................... 38

Figure 3.1: Map of the study area and populations........................................................................ 78

Figure 3.2: a) Probability of data as a function of K (the number of clusters) in STRUCTURE analysis suggest K=8 is the optimal number of clusters. b) Changes of $\Delta K$ as a function of K (the number of clusters) suggest K=4 as the optimal number of clusters according to Evanno method ...................................................................................................................... 79
Figure 3.3: Assignment of individuals to genetic clusters using a) Bayesian clustering program STRUCTURE and b) Discriminant analysis of principal components (DAPC) method at optimal genetic clusters (K=8). Both methods show very similar pattern of individual assignments to genetic clusters................................................................. 80

Figure 3.4: Discriminant analysis of principal components (DAPC) scatter plot with only Alouette and Coquitlam individuals. ...................................................................................................................... 81

Figure 3.5: Bar plot showing the assignment of individuals (sea-run and resident) from the Coquitlam reservoir into two genetic clusters (K=2) using DAPC method show a clear distinction in their genetic makeup between sea-run and resident individuals................................. 82

Figure 3.6: Results of IM analysis showing estimated population sizes, and population splitting time. Time is represented in the vertical axis. Widths of boxes are representative of current and ancestral population size estimates, and horizontal line represent population splitting time (95% highest posterior density (HPD) in dashed horizontal lines). .................................................. 83

Figure 3.7: Difference between historical and contemporary gene flow (NmH – NmC) estimates based on Migrate-n and BayesAss analysis. Positive values indicate higher historical gene flow relative to contemporary gene flow. ................................................................. 84

Figure 4.1: Location of the study area in Northwest Territories, Canada. Northern Canada (north of 60°N latitude) contains the three Canadian territories: Nunavut, Northwest Territories, and Yukon................................................................. 114

Figure 4.2: Occurrence of each fish species in 53 lakes and ponds in the Barrenlands region of Northwest Territories, Canada. ARGR: Arctic grayling; BURB: Burbot; CISC: Cisco; LKCH: Lake chub; LKTR: Lake trout; LKWH: Lake whitefish; LNSC: Longnose sucker; RNWH: Round whitefish; SLSC: Slimy sculpin................................................................. 115

Figure 4.3: Relationship between lake area and species richness for Barrenlands lakes. Dark circles are species richness estimates for fish-out lakes and open circles are species richness estimates for monitored lakes. Regression line shows the overall trend for all lakes. .............. 116
Figure 4.4: Relationship between species richness and lake area for lakes in different regions in North America. There was no statistical difference in slopes among regions but intercepts were significantly different. The common slope was plotted with the region specific intercept because slopes were not significantly different. Positioning of the symbol indicates the average lake size for that region. Note that regression lines for Black-Hollow region and Manitoulin Island coincide. ................................. 117

Figure 4.5: Total fish biomass distributions for Barrenlands lakes (light grey) and Ontario lakes (dark grey) on a shared axis. The two distributions were very marginally different ($p = 0.08$; Wilcoxon rank-sum test). ............................................. 118

Figure 4.6: Relationship between fish biomass, lake area, mean annual air temperature, and species richness. a) Log$_{10}$ transformed total fish biomass as a function of Log$_{10}$ transformed lake area. Dark circles represent Barrenlands lakes. b) Log$_{10}$ transformed biomass per area (in kg/ha) as a function of mean annual air temperature with numerical symbols representing the species richness. Barrenlands lakes are located around -10 °C. ............................................. 119

Figure 4.7: a) Mean percentage abundance of the dominant species for lakes with more than one species. b) Mean percentage biomass of the dominant species for lakes with more than one species. Dark circles indicate the mean percentage biomass of the dominant species from Ontario lakes (calculated from Kelso & Johnson 1991). Error bars represent the standard error of the mean................................................................. 120

Figure 4.8: Lake trout biomass in Canadian lakes. There was no a significant linear relationship between lake trout biomass and latitude ($R^2 = 0.01$, $p = 0.44$). Dark circles represent lake trout biomass in Barrenlands lakes................................................................. 121
List of Appendices

Table A 2.1: Mean square error (MSE) and relative mean square error (RMSE) for Migrate-n and BayesAss estimates for all combinations of gene flow change, population size and time since the change in migration....................................................................................................................... 39

Table A 4.1: Summary of data used for analysis of fish biomass......................................................... 122
1.1 General background

Conservation biology is an applied field of biology that developed to address concerns about the loss of biodiversity; as such, the primary goal of conservation biology is to study biodiversity to develop approaches to prevent biodiversity loss (Soule 1985). Ecology and evolutionary biology are cornerstones of conservation biology and conservation biologists apply ecological and evolutionary theory to better understand biodiversity and devise plans to maintain it.

Biodiversity is typically described at three levels of biological organization: 1) genetic diversity at the population or species level; 2) species diversity at the community level; and, 3) community diversity at the ecosystem level. Initially, conservation efforts focused on protecting taxonomically recognized species (Caughley 1994); but federal conservation legislations such as the Endangered Species Act in the United States and the Species at Risk Act of 2002 in Canada allow listing of biodiversity units below the species level today. Although there is some debate as to which level of biological organization (i.e. population, species, or community) is the most effective conservation target, a “one-size-fits-all” solution is unlikely to be effective as every conservation issue has its own unique circumstances.

1.2 Scale considerations in conservation

Scale is an important consideration in biology, which spans the range from the study of biological molecules (molecular biology), to the study of how biotic and abiotic systems interact at large spatial scales (ecosystem biology). Conservation biologists must also consider scale because the scales of conservation concerns vary. For example, a conservation target might be a
single population, several populations of a particular species, the species as a whole, a
community with many species, or multiple communities. Similarly, threats to biodiversity can
also vary in spatial scale. Some threats, such as climate change or large-scale development
projects (e.g., resource exploitation and human settlements) are broad, and can potentially affect
multiple communities composed of many species. In contrast, some threats are localized, such as
a targeted fishery in a specific area that only affects one or a few populations of a single species.
Thus, the conservation target can vary at both the level of biological organization and the spatial
scale. Therefore, mitigating threats to biodiversity requires working at the appropriate scale to
address the issue at hand.

1.3 Habitat changes in space and time

Habitat loss and fragmentation through agriculture, deforestation, natural resource extraction
(e.g. mining), and other activities related to the expansion of the human population have changed
the global landscape drastically in the last 500 years (Noss et al. 1995). As a result, habitat loss
and fragmentation are considered to be the primary cause of biodiversity loss at the regional and
global scales (Wilcove et al. 1998; Balmford et al. 2005; Venter et al. 2006). Surveys in the
United States have documented over 98% declines in old growth forests, grassland, and savanna
habitat (Noss et al. 1995). Not only has the landscape changed, it has changed at an increasingly
rapid pace, paralleling the rapid increase in the human population in the last two centuries. Some
noteworthy examples of recent habitat loss include: greater than 50% decline in the pre-industrial
range sizes for 72% of global mammal species studied by Ceballos & Ehrlich (2002); declines in
range sizes of British birds (54%), vascular plants (28%), and butterflies (71%) between 1960-
1990 (Thomas et al. 2004); and, an annual loss of tropical forest cover at a rate of 0.49%
between 1990-2010 (Achard et al. 2014).
Chapter 1: Introduction

Although temporal analysis of global habitat fragmentation is unavailable (likely because no single measure can capture the complexity of fragmentation and because different researchers use different methodologies), fragmentation is thought to have increased within the last century due to proliferation of roads, dams, and other human infrastructure. A few studies at local or regional scales have shown that fragmentation has increased over time. For example, in the Brazilian Amazon, the number of forest patches per 100 ha area has tripled and distance between patches has doubled from 1984 to 2002 (Ferraz et al. 2005). A recent study by Ahmed et al. (2013) found that the road network in the Brazilian Amazon grew by almost 17,000 km per year between 2004 and 2007. In Baden-Württemberg, Germany, the effective mesh-size (a measure of connectivity between two random points in the landscape) has decreased by about 40% between 1930 and 2004 due to increases in transportation infrastructure (Jaeger et al. 2007). These findings all indicate significant increases in the levels of fragmentation in recent years.

Habitat loss and degradation (including habitat fragmentation), overexploitation, water pollution, climate change, and invasive species are the main threats to aquatic biodiversity (Wilcove et al. 1998; Dudgeon et al. 2006; Venter et al. 2006; Heino et al. 2009). Habitat change, which includes habitat loss and fragmentation, and loss of suitable habitat through pollution and climate change, is likely to continue to be the key threat to aquatic biodiversity in the future. The estimated loss of aquatic biodiversity between 1970 and 2000 (especially for freshwater systems) exceeds that of terrestrial ecosystems (Loh et al. 2005). Unfortunately, quantifying aquatic habitat change is a challenge. Surprisingly, the global extent and distribution estimates of wetlands, rivers, and lakes is poorly known, hence, it has been difficult to quantify the extent of recent losses with reasonable accuracy (Revenga et al. 2005). However, it is estimated that about 52% of wetlands in the US have been lost since European settlement. Between 1986 and 1997,
Chapter 1: Introduction

US wetlands have decreased at a rate of 0.06% per year (Revenga et al. 2005). Decreases in the quality of freshwater habitat through pollution and degradation, flow changes, and fragmentation are also difficult to quantify. However, one cause of fragmentation, dams, can be quantified. Globally, a total of 26,708 large dams (>15m high) came into operation between 1940-2000 (International Commission on Large Dams), and today, approximately 60% of large river systems are fragmented by dams (Nilsson et al. 2005). Therefore, it is not unreasonable to infer that aquatic habitat degradation was significant in the last century.

1.4 Status of Canadian fishes

In Canada, habitat loss and degradation are the primary threats for freshwater fishes, and overexploitation is the primary threat to marine fishes (Dextrase & Mandrak 2006; Venter et al. 2006). According to Venter et al. (2006), 68% of at-risk freshwater fish were affected by habitat loss and 51% of freshwater fishes were affected by pollution, which ultimately leads to loss of fish habitat. According to Dextrase & Mandrak (2006), 73% of at-risk freshwater fishes were affected by habitat loss or degradation. Looking forward, climate change has been identified as a potential key threat to Canadian fishes, especially in northern Canada (Reist et al. 2006). Invasive species are also expected to be a major threat for freshwater fishes (Dextrase & Mandrak 2006). Ultimately, multiple threats will affect Canadian fish populations and how these threats are mitigated will determine the long-term persistence of fishes in Canada. Further concerns have been raised about the conservation of Canadian fishes in light of recent changes to the Fisheries Act, which, it is thought, will reduce the protection previously afforded to aquatic organisms because of the relaxation in environmental scrutiny for development projects (Hutchings & Post 2013; Gantner 2014). The changes to the Fisheries Act limit protection to only fishes that are part of an existing fishery. It is thought that this will accelerate degradation of
Chapter 1: Introduction

Aquatic habitat through resource and infrastructure development in remote places, as remote areas are unlikely to have active fisheries (Hutchings & Post 2013). This change in the Act is especially concerning as a recent study by Chu et al. (2015) suggests that cumulative environmental and anthropogenic stressors are expanding northwards into the more remote areas in Canada. As resource development and human infrastructure move northwards, further aquatic habitat loss and degradation are likely to follow.

1.5 Thesis objectives and motivation

In light of recent, rapid changes to habitat, wildlife populations, and natural communities worldwide, this thesis examines effects of such changes from a conservation perspective. The goals of the thesis are to address gaps in our understanding of conservation at the population and community levels. I focused on issues that are important for making well-informed decisions based on the identification of population structure and connectivity, assessing risks, and conserving biodiversity at the appropriate scales. To address questions related to biodiversity conservation, I investigated the effects of habitat change in space and time, and at different scales of biological organization. I used simulated and empirical data to address these questions. I begin with a simulation study investigating the effects of recent population connectivity changes on genetic methods for inferring population connectivity. This is followed by two chapters, focused on empirical, biodiversity data for Canadian fishes, a valuable, renewable resource for Canadians. Insights from this thesis can contribute to sustainable management and conservation of biodiversity in Canada and potentially world-wide.

Chapter 2 explores the theme of how the timing of population connectivity change can affect our inferences of population connectivity in spatially structured species of conservation concern. Currently, molecular ecologists typically use coalescent-based and disequilibrium-based methods
Chapter 1: Introduction

to infer genetic connectivity among populations. However, recent changes to population
connectivity could impact the accuracy of those estimates. I assessed the effectiveness of those
methods under a range of scenarios using simulated genetic data. The driving question in this
chapter is: how well can we estimate genetic connectivity among populations when connectivity
has changed in an ecologically relevant time frame? This is an important question because
human population growth and infrastructure have changed the global landscape drastically over
the last 500 years, and especially over the last 100 years. The results of my investigation should
be relevant for any species whose populations may have undergone connectivity changes in
recent times and should thus have broad applicability.

In Chapter 3, I use empirical data, collected on populations at a relatively small scale of spatial
and biological organization, to investigate the effects of habitat fragmentation on the ecology and
evolution of eleven Sockeye salmon (*Oncorhynchus nerka*) populations in the Lower Fraser
River region. Specifically, I investigated the effects of two hydroelectric dams on population
genetic structure and the connectivity of Sockeye salmon populations in this region. It was
believed that the anadromous form was lost from both reservoirs when the dams were built, but
recent observations of downstream migration of salmon during an experimental water-release
program warranted re-examination of this hypothesis. I compared and contrasted historical and
contemporary genetic connectivity patterns to infer possible changes that might be related to the
dams and other recent habitat changes. In the process, I also investigated the genetic distinction
and evolution of two life history forms of Sockeye salmon living in reservoirs created by dams.

In Chapter 4, I explore phenomena related to fish biodiversity at larger scales of spatial and
biological organization. I characterized and tested hypotheses at the community level by
examining fish species diversity and biomass in lakes in the Northwest Territories that are
Chapter 1: Introduction

affected by increasing resource exploitation. Increasingly, northern Canada is experiencing significant change because of resource exploration and extraction (Chu et al., 2015) but our understanding of the ecology of organisms in this region is relatively limited. It is important to characterize and understand biological resources in these systems to bring more attention to the biodiversity that may be lost to increasing habitat destruction in northern Canada. Northern Canada is generally considered to be very poor in biodiversity (e.g., species diversity) and fish resources (quantity of fish that can be exploited), making regulatory approval for development projects relatively easy. Therefore, scientific studies on biodiversity in this region are needed to provide more information to regulators and highlight the value of fishes in northern Canada. I assessed fish communities impacted by mining, which is causing habitat destruction at a large spatial scale and could affect many aquatic species. I focused on whole-lake communities in the Barrenlands region of Northwest Territories. Here, whole lakes are being drained to create mining pits and waste containment facilities. I utilized data collected by Fisheries and Oceans Canada through environmental consulting firms to quantify fish communities and resources in this system and compare them with lakes from southern Canada. This chapter has been published in the Canadian Journal of Fisheries and Aquatic Sciences.

In summary, in this thesis, I investigate the effects of habitat change in space and time on Canadian fish diversity. I examine conservation issues at different spatial and biological scales, which are targets for conservation action. Using a combination of simulation-based and empirical approaches, I address gaps in conservation biology in the hope of ultimately helping conservation practitioners to conserve biodiversity.
1.6 Literature cited


Chapter 1: Introduction


Chapter 1: Introduction


Chapter 2

Genetic connectivity in a changing world: How accurately can we estimate current genetic connectivity when population connectivity has changed in recent times?

2.1 Abstract

An accurate understanding of population connectivity is important for ecologists, evolutionary biologists and conservation biologists because migration plays an important role in population dynamics, microevolution, and assessment of extirpation risk and chance of population rescue. Genetic methods are increasingly used to infer connectivity among populations as advances in technology have made it cost effective and advantageous compared to ecological methods. However, how well genetic connectivity inference methods perform when connectivity has declined substantially in recent times has not been well investigated. Given drastic changes to wildlife population connectivity within the last century, errors or biases in connectivity estimates could be especially important in a conservation context. Using simulated genetic data, we investigate the accuracy and bias of connectivity estimates from two commonly used genetic inference methods when connectivity has declined recently. We focus on the timing of connectivity change and the magnitude of change on estimates of coalescent-based and disequilibrium-based inference methods. Contrary to expectations, the coalescent-based method (Migrate-n) is sensitive to recent changes and provides better estimates of current connectivity than the disequilibrium-based inference method (BayesAss). Both methods generally overestimate current connectivity when connectivity has been reduced recently. Our results also show that, in situations where historically high connectivity has been reduced recently, coalescent-based estimates should not be thought of as reflecting long-term migration rates,
because they underestimate the historical rate. Overall, our results highlight the problem with using coalescent estimates as historical rates and disequilibrium estimates as contemporary rates when connectivity has changed in recent times. Conservation biologists should be cautious when they use these estimates and we provide some general guidance based on our experience.

2.2 Introduction

In today’s dynamic world, a substantial number of wildlife populations are experiencing rapid declines, many of which are driven by human activities. Human population expansion and industrialization have resulted in an unprecedented level of global change from habitat degradation and loss, to climate change that is having broad effects across landscapes. Loss of habitat, land-use change, and human infrastructure (e.g. roads, dams) can result in a reduction in connectivity among populations. Connectivity can play an important role, via both ecological and evolutionary effects, on genetic diversity, adaptation, and ultimately, the long term persistence of wildlife populations (Frankham 2005; Broquet & Petit 2009). Isolation of populations results in a reduction in gene flow and an increase in genetic drift, resulting in the loss of genetic diversity over time (Nei 1973). In turn, a loss of genetic diversity can lead to inbreeding depression, a reduction in the average fitness of the population, and a reduction in populations’ ability to adapt to novel pressures (Keller & Waller 2002; Reed & Frankham 2003). Reduced population connectivity also reduces the chance of population recovery through immigration. Therefore, getting an accurate picture of population connectivity is of paramount importance to ecologists, evolutionary biologists and conservation practitioners.

Estimating the level of connectivity among populations of at-risk species, which have become partially isolated because of recent habitat loss or fragmentation, is of great interest to population
ecologists who monitor these populations, evolutionary geneticists who are concerned with the loss of genetic diversity, and conservation managers who are charged with creating action plans to prevent extirpation and, over the long term, species extinction. However, estimating connectivity among populations is a challenging task. Increasingly, molecular methods, rather than ecological methods, are used to estimate migration because of the difficulties (e.g., tracking is labour intensive) and limitations of ecological methods (e.g., the probability of detecting rare, long-distance movements can be very low at large spatial scales - (Faubet et al. 2007). There have been recent advances in genotyping and sequencing technologies and, consequently, improved genetic methods for estimating movement have been the focus of intensive research (Broquet & Petit 2009; Dyer et al. 2010). In particular, determining changes in population genetic connectivity over time is an area of research that has important practical implications for biodiversity conservation (Kinnison & Hairston Jr 2007; Faubet & Gaggiotti 2008).

Two types of methods for estimating connectivity have been used extensively in the conservation and molecular ecology literature: (i) coalescent-based estimators are used for estimating historical, long-term genetic connectivity; and, (ii) disequilibrium-based methods are used to estimate current (i.e. recent) levels of genetic connectivity among populations. Both of these estimators have been used to make inferences about population connectivity for biodiversity conservation purposes on numerous occasions (e.g., (Howes et al. 2009; Dubey & Shine 2010; Welch et al. 2012; Huey et al. 2013; Sharma et al. 2013; Klimova et al. 2014). To test hypotheses related to human-induced changes in genetic connectivity through habitat fragmentation, some authors have contrasted long-term and contemporary estimates of migration rates (Chiucchi & Gibbs 2010; Apodaca et al. 2012; Sharma et al. 2013; Blakney et al. 2014). For example, Chiucchi and Gibbs (2010) found similar historical and contemporary migration rates in
rattlesnake populations and concluded that recent fragmentation had had minimal effect on their genetics.

To our knowledge, no one has rigorously examined the effects of recent changes in population connectivity on inferences about connectivity derived from these methods by using information about connected populations with known migration rates. Such analyses would allow us to determine whether inference methods can accurately estimate contemporary connectivity among populations when human-induced habitat fragmentation has caused a recent (e.g., in the last 50 generations of a species) decline in connectivity. Given that recent human activities are reducing population connectivity for many species, and estimates of connectivity play an important role in management and recovery strategies for at-risk species, it is important to understand the accuracy of inferences derived from these methods.

Here we use simulations of populations with known rates of migration among them to determine the sensitivity of coalescent and disequilibrium-based estimators of migration rate to recent changes in genetic connectivity. We evaluate Migrate-n and BayesAss: two methods frequently used to estimate historical and contemporary migration rates, respectively. Migrate-n uses coalescent theory (Kingman 1982) and molecular data (e.g., DNA sequence or microsatellite data) to estimate historical population size and migration rates. In general, migration rates estimated by Migrate-n represent historical, long-term migration spanning, roughly, the last 4Nₑ generations (where Nₑ is the effective population size), assuming long-term equilibrium. In contrast, BayesAss uses transient disequilibrium observed in multilocus genotypes of migrants to estimate contemporary migration rate (Wilson & Rannala 2003). Therefore, migration rates estimated with BayesAss are thought to be indicative of the last 1-3 generations (Howes et al.)
Chapter 2: Genetic connectivity in a changing world

2009; Huey et al. 2013). We are particularly interested in the effects of recent (e.g., within the last 100 generations of a species) declines in gene flow because they are thought to be common, and because of its potential implications for estimating extirpation risk.

Based on the underlying theoretical frameworks used by these two methods, we predict that Migrate-n will be less accurate than BayesAss at estimating current population connectivity because coalescent approaches (like Migrate-n) assume stable dynamics in migration over the last ~4Ne generations (Kuhner 2009). Since most relevant “recent” declines in migration have likely occurred within the last 4Ne generations (unless Ne is extremely small), we expect Migrate-n to overestimate the current level of genetic connectivity and be closer to the historical, higher level of connectivity. Furthermore, we expect the accuracy of Migrate-n to decline as the time since the change in the migration rate decreases.

We predicted that BayesAss should have better accuracy under a variety conditions than Migrate-n because it does not assume equilibrium conditions. In addition, because BayesAss estimates migration within the last 1-3 generations, it should provide more accurate estimates of contemporary connectivity and should not be affected by migration changes that occurred more than three generations before sampling. In summary, we expect the timing of changes in migration to affect the accuracy of Migrate-n estimates of genetic connectivity but not of BayesAss’ estimates. We also predicted that the magnitude in the change of migration would have an effect on the accuracy of estimates. For Migrate-n, we expected a large decline in recent migration to have a substantially larger effect on estimate accuracy than a small decline in migration. In contrast, for BayesAss, the accuracy should not be affected if the change in migration occurred more than 3 generations before present.
We tested these ideas by simulating population genetic data for four connected populations that exchanged migrants at a high rate historically but at a low rate in recent times. The genetic data were sampled at the end of the simulation period to estimate population connectivity using Migrate-n and BayesAss methods. These estimates were compared against the contemporary migration rate used in the simulation to investigate the accuracy and bias of estimates from both methods.

2.3 Methods
We focused on the effects of (i) the time since a change in migration \( (t) \), and (ii) the magnitude of gene flow change \( (\Delta N_m) \) on estimates of migration rate using genetic data. The time since a change in migration \( (t) \) was quantified in terms of generations (i.e., number of generations between the change in migration and the sampling of genetic data). The magnitude of the change in gene flow \( (\Delta N_m) \) was defined as the difference in the number of migrants \( (N_m) \) before and after the change, where \( N \) is the population size and \( m \) is the immigration rate into the focal population. We also conducted simulations with different population sizes \( (N) \) to determine whether population size has an effect on the accuracy of the estimates. The \( m \) was adjusted when \( N \) changed, to keep the gene flow level \( (N_m) \) consistent. To summarize, we simulated population genetic data (see details below) and varied the following parameters: (i) the timing of the change in the migration rate \( (t) \); (ii) the magnitude of migration change \( (\Delta N_m) \); and, (iii) population size \( (N) \). These simulated genetic data were then used as input for the Migrate-n and BayesAss programs to estimate migration parameters and the estimates were compared with the known migration parameter used for simulating the data.
2.3.1 Simulations of genotype data

We simulated microsatellite genotype data using the program EASYPOP 2.0.1 (Balloux 2001), which simulates population genetic data evolving forward in time. The goal was to simulate populations that were connected by migration for most of their histories at a rate of one migrant per generation (Nm = 1), but that then experienced a recent decline in connectivity. A road constructed 50 years before present that completely cut-off migration between previously connected populations would be an extreme example of this (Figure 1). Other situations, such as habitat conversion that reduces effective migration, would result in decreased migration relative to historical levels. For each scenario, we simulated a set of four populations with equal size (N), each with randomly mating individuals within a population, and each exchanging migrants with the other three populations with the probability \( m \) (per generation) according to Wright’s Island model (Figure 1a). All populations were evolved forward in time for a total of 5000 generations, but the change in migration occurred at different time points before the end of the simulation; to investigate the effect of time since the change in migration, a change in the migration rate was imposed \( t \) generations before the end of the simulation (where \( t = 500, 200, 100, 50, 20, 10 \) and 5 generations before the end of the simulation). To simulate different magnitudes of gene flow declines, the gene flow was initially set at one migrant per generation (Nm before = 1), and was reduced to one of four possible levels after the change (Nm after = 0, 0.25, 0.5, 0.75). This resulted in four classes of levels of change in connectivity: (i) Extreme (Nm = 1 to 0, \( \Delta N_m = 1 \)); ii) Severe (Nm = 1 to 0.25, \( \Delta N_m = 0.75 \)); (iii) Intermediate (Nm = 1 to 0.5, \( \Delta N_m = 0.5 \)); and (iv) Moderate (Nm = 1 to 0.75; \( \Delta N_m = 0.25 \)). At the end of each replicate simulation (5000 total generations), 40 individuals from each population were sampled for genetic analysis (i.e., 160 individuals per data set). All genetic loci had the same mutation dynamics: 30 independent
Chapter 2: Genetic connectivity in a changing world

microsatellite loci mutating according to a single-step mutation model with a mutation rate of \( \mu = 5 \times 10^{-4} \). Genotypes for individuals in the initial generation were randomly drawn from all possible allelic states. We examined the effects of population size on the migration estimates by performing simulations at four different population sizes (N=100, 200, 1000, 2000). At each population size, \( m \) was adjusted to get the desired level of gene flow (Nm). In total, 20 replicate data sets were simulated for each combination of variables (\( t \times \Delta Nm \times N \)) for a total of 112 variable combinations and 2240 data sets.

2.3.2 Analyses

Genetic analyses to estimate migration rates were conducted using Migrate-n 3.2.16 (Beerli 2006) and BayesAss 1.3 (Wilson & Rannala 2003). Migrate-n analyses were conducted using Bayesian inference with slice sampling. The following parameters were used for Markov Chain Monte Carlo runs in the program: Brownian motion with constant mutation rate, recording 15000 steps at 75 step increments (total steps 1,125,000) and the first 15,000 genealogies were discarded as burn-in. We used the UPGMA tree as the starting genealogy and used static heating with six temperatures as specified by the # command in the program for even spacing (temperatures were 1.00, 1.25, 1.67, 2.50, 5.00, 1000000.0). Posterior distribution, autocorrelation, and effective sample size (ESS) were checked to judge convergence. Please refer to the Migrate-n program manual (Beerli 2012) for a detailed description of parameters and commands.

BayesAss analyses were conducted with MCMC runs of \( 10 \times 10^6 \) iterations with a sampling frequency of 2000 and discarding the first \( 2 \times 10^6 \) iterations as burnin. The delta values were initially set to 0.15 and this provided adequate level of mixing (40%-60%). We performed four
independent runs with different seed numbers for each replicate and the run with the minimum
Bayesian deviance was used in the migration estimate analyses (see Faubet et al. 2007 and
Spiegelhalter et al. 2002 for more detail).

2.3.3 Evaluations of accuracy and bias

Accuracy of the estimated migration rates were evaluated using relative mean square error
(RMSE; Equation 1), except in scenarios where \( N_{m_{after}} = 0 \), in which case mean square error
(MSE; Equation 2) was used instead. The \( m^* \) is the contemporary migration rate that was used to
simulate data.

(1)

\[
RMSE (m) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{m_i - m^*}{m^*} \right)^2
\]

(2)

\[
MSE (m) = \frac{1}{n} \sum_{i=1}^{n} (m_i - m^*)^2
\]

Bias in the estimated migration rates were evaluated using relative bias (RBias; Equation 3),
except in scenarios where \( N_{m_{after}} = 0 \); in these cases, bias (Bias; Equation 4) was used instead.

Using relative mean square error (RMSE) and relative bias (RBias) allowed comparisons among
estimates for varying migration rates because they standardize the estimates by dividing by the
migration rate used for simulation. We observed that the trends observed within a set of
simulations using relative (e.g., RMSE) or absolute (e.g., MSE) quantities were identical and,
Chapter 2: Genetic connectivity in a changing world

therefore, here, we present RMSE except when MSE is required because RMSE cannot be calculated (i.e., when $N_{m_{after}} = 0$)

\[
R_{Bias} (m) = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{m_i - m^*}{m^*} \right)
\]

(3)

\[
Bias (m) = \frac{1}{n} \sum_{i=1}^{n} (m_i - m^*)
\]

(4)

To determine how the timing of the change in gene flow affected accuracy, we calculated the ratio of mean square error at 5 generations to mean square error for 500 generations ($MSE_5/MSE_{500}$). When this ratio is substantially above 1, this indicates that the recent gene flow change resulted in higher error than when the change was historical. Accuracy and bias were calculated using R version 2.14.2 (R Development Core Team 2012).

2.4 Results

Both the overall accuracy (evaluated using relative mean square error or mean square error) and the bias in estimates were influenced by all of the parameters that we investigated: the timing of the gene flow change, the magnitude of gene flow change, and population size. However, the extent to which these factors affected accuracy and bias differed for the two programs.

Generally, estimates from both methods were higher than the actual contemporary migration rate used to simulate the data. For the most part, Migrate-n migration rate estimates were closer to the actual migration rate than BayesAss estimates, which substantially overestimated the current
migration rate. BayesAss estimates had higher relative mean square error (or mean square error) and positive bias compared to Migrate-n estimates. Below, we describe in more detail the observed trends in accuracy and bias with respect to the timing of change in connectivity, the magnitude of change, and population size.

2.4.1 Effect of the timing of gene flow change

In general, the accuracy of the estimated migration rate declined as the time since the change in connectivity was reduced (Figure 3, 4, 5; Appendix Table 1). In other words, migration rate error is higher when the change was relatively recent compared to when it was historical (i.e., a change 500 generations ago provided a more accurate estimate than a change 5 generations ago). For both programs, the relative mean square error (RMSE) tended to decrease as the time since the connectivity change increased, but this trend was stronger for Migrate-n estimates compared to BayesAss estimates and this general trend was most apparent when gene flow changed from Nm of 1 to 0 (Figure 3); here, the Migrate-n MSE was generally highest when change occurred 5 generations before sampling and lowest when change occurred 500 generations before sampling. For BayesAss, at smaller population sizes (100 and 200), MSE tended to decrease with increasing time since the change but, after 100 generations, the error seemed to be relatively stable (Figure 3). In contrast, for BayesAss, at large population sizes (e.g., 1000, 2000), MSE appeared to be independent of the time since the change. The ratio of MSEs for generations 5 and 500 (MSE$_5$/MSE$_{500}$) shows the sensitivity of these methods to the timing of the connectivity change. Migrate-n ratios deviate more from 1.0 than those for BayesAss, indicating that Migrate-n estimates are more sensitive to the timing of the gene flow change than BayesAss estimates (Table 1).
Chapter 2: Genetic connectivity in a changing world

2.4.2 Effect of the magnitude of gene flow change

For both Migrate-n and BayesAss, the accuracy declined (i.e., RMSE increased) as the magnitude of gene flow change increased. For example, with a population size of 200 (N=200) and a gene flow change five generations before sampling, the BayesAss RMSE was 1.28 when gene flow changed from Nm of 1 to 0.75 (25% decline), 1.92 when gene flow changed from Nm of 1 to 0.5 (50% decline), and 7.10 when gene flow changed from Nm of 1 to 0.25 (75% decline) (Figure 4). This pattern of increasing error with increasing magnitude of gene flow change appears to be stronger for BayesAss than Migrate-n (Figure 4; Appendix Table 1). As for RMSE, relative bias (RBias) also increased with increasing magnitude of gene flow change. For both methods, there was a positive bias when the magnitude of the change in gene flow was high (e.g. Nm 1 to 0.25), and this positive bias declined and became slightly negative as the magnitude of the change in gene flow (ΔNm) decreased (e.g. Nm 1 to 0.75).

2.4.3 Effect of population size

The size of the simulated populations had an effect on the accuracy and bias measures of the migration rate estimate. In general, as the size of the population increased, the relative mean square error (RMSE) increased (comparing RMSE across population sizes for a particular level of gene flow change (ΔNm) and for a particular time since the gene flow change (t)). This trend was consistent across all the simulated scenarios. Similarly, the relative bias in the migration rate estimate tended to increase (i.e. became more positive) as the size of the simulated population increased. For example, in simulations when the gene flow declined from Nm of 1 to 0.75, the relative bias was negative for a population size of 100 for both Migrate-n and BayesAss, but the bias was positive for both programs when population sizes were 1000 and 2000.
Chapter 2: Genetic connectivity in a changing world

2.4.4 Accuracy and bias when time is scaled to population size

By scaling the time since the connectivity change from number of generations to generations relative to the size of the population (i.e., by dividing by N), we can summarize across simulations with different population sizes and identical gene flow changes. When the migration timing change was scaled to population size, the results showed that both relative mean square error and relative bias increased as the time since the migration rate change decreased (Figure 5). This trend was observed for both BayesAss and Migrate-n estimates, and was consistent for different magnitudes of change in migration. Furthermore, as for previous results, BayesAss estimates consistently had higher relative mean square errors than Migrate-n.

2.5 Discussion

In the empirical literature, estimates of migration among populations from coalescent-based methods such as Migrate-n are interpreted as historical, long-term rates and disequilibrium-based BayesAss estimates are interpreted as migration rates in very recent generations. The primary focus of this study was to determine the accuracy of these methods when there was a relatively recent change in connectivity. To our knowledge, this is the first study to investigate this issue, which has important implications for biodiversity conservation because of recent, human-induced declines in connectivity of many wildlife populations. Our results show that recent changes in connectivity can indeed cause both methods to overestimate current connectivity. For both methods, the timing of the connectivity change affected migration estimates; we expected this for coalescent-based Migrate-n estimates, but we did not expect to see such an effect with BayesAss. We also found that the magnitude of the change in connectivity had an influence on estimates, as did population size. Below, we discuss our key findings, which will be points of
interest for molecular ecologists and conservation biologists who conduct empirical studies in a conservation context.

2.5.1 Historical connectivity

In situations where an estimate of historical, long-term connectivity is the parameter of interest, coalescent-based methods are often used to quantify gene flow. Migrate-n is among the most commonly used programs to estimate this but other programs such as LAMARC (Kuhner 2006) and IMa2 (Hey 2010) are also used, depending on the research question and context. Our results suggest that, in situations where genetic connectivity has changed substantially in recent times, the estimates by coalescent-based methods (Migrate-n in particular) do not provide as accurate an estimate of the historical connectivity as expected by conservation biologists and molecular ecologists. Our review of the empirical literature suggests that most conservation biologists and molecular ecologists interpret coalescent-based gene flow estimates as long-term, historical, average gene flow over the last $4N_e$ generations (Chiucchi & Gibbs 2010; Hsieh et al. 2013; Sharma et al. 2013; Blakney et al. 2014).

The problem with interpreting coalescent-based estimates as long-term, historical connectivity is apparent from the results of our simulations where recent connectivity was relatively low compared to historical levels. As our results on the accuracy and bias of estimates at different time points show, estimates were affected by the time since the change in migration (Figure 3). In our simulations where historical connectivity was higher than recent connectivity for most of the simulation period, Migrate-n estimates underestimated the historical level of genetic connectivity. This shows that the recent level of migration can have a substantial influence on the coalescent estimate, resulting in an underestimate of the historical rate if the recent rate is lower than the historical rate.
2.5.2 Recent connectivity

Although we expected BayesAss to be more accurate than Migrate-n for estimating current migration rates when connectivity had changed in recent times, our results show that, in fact, in most situations, Migrate-n estimates were closer to actual, recent levels of connectivity than BayesAss estimates. Faubet et al. (2007) tested the accuracy of BayesAss estimates of migration rate at different levels of migration and $F_{st}$ values but, unlike our study where migration rate declined, they kept the rate constant. They concluded that the estimates are reasonably accurate if the migration rate is low ($m=0.01$) and genetic differentiation among populations is high ($F_{st} \geq 0.1$); however, Faubet et al. (2007) noted that the accuracy was poor when migration rate was high ($m \geq 0.05$), and the method is very sensitive to allele frequency changes. In our simulations, the highest initial migration rate was 0.01 and the final rate (i.e. after the decline) rate was always less than 0.01, therefore, our migration rates were well below the problematic levels identified by Faubet et al. (2007). However, changes in allele frequencies due to effects of genetic drift (caused by finite population sizes and changes in the migration rate) may have caused problems for the BayesAss inference model. In our experience, BayesAss appears to have a lower detection threshold for $m$ at about 0.0025. Combining our findings with those of Faubet et al. (2007), we conclude that BayesAss is likely to be poor at estimating migration rate for large populations ($N_e \geq 1000$) because either the actual $m$ is likely to be below the lower limit of estimation ($m \leq 0.0025$) or the gene flow ($N_e m$) between populations is high enough (because $N_e$ is high) to cause problems for detection of immigrants ($F_{st}$ between populations is too low). For example, a population with an effective size of 1000 ($N_e =1000$) and a migration rate of 0.0025 (lowest detection threshold for BayesAss in our experience) would have an $F_{st} \approx 0.09$. If the migration rate is higher than 0.0025, $F_{st}$ will be well below 0.1, which also reduce the accuracy
of migration rate estimates according to Faubet at al. (2007). Therefore, BayesAss would be poor at estimating the migration rate for large populations because the actual migration would be below its estimation threshold or because $F_{st}$ among populations is too low to accurately estimate migration rate. Furthermore, our results also showed that BayesAss estimates are influenced by the timing of the change in connectivity (Figure 3). Consequently, in situations where population connectivity was high historically but has declined recently, BayesAss would overestimate the current migration rate.

2.5.3 Conservation implications

Our findings have significant implications for biodiversity conservation because it is important to accurately identify and quantify patterns and degree of population connectivity for multiple reasons. For example, it is relevant when the probability of natural rescue by migrants from nearby populations (i.e., the rescue effect) needs to be determined. It is also important for assessing long-term maintenance of genetic diversity and population persistence. We found that, in general, the estimates of genetic connectivity and gene flow inferred by Migrate-n and BayesAss are likely to overestimate the current level of connectivity; this is particularly true if one is using BayesAss. Therefore, molecular ecologists and conservation practitioners should be cautious about interpreting the estimates of connectivity if there is any evidence to suggest that connectivity might have declined in recent times. This could be straightforward knowledge in some situations (e.g., a road or dam was constructed 50 years ago). However, it could be less clear in situations where more subtle changes in land use or habitat type might have reduced effective migration among populations. Overall, we suggest that estimates produced by BayesAss and Migrate-n be interpreted as upper limits of connectivity with the recognition that the actual level of connectivity may be considerably lower.
Chapter 2: Genetic connectivity in a changing world

Our results also point to further concerns in interpreting these estimates; some authors have contrasted coalescent estimates of connectivity (interpreted as historical connectivity) against disequilibrium based estimates of connectivity (interpreted as contemporary connectivity) to investigate changes in population connectivity over time. For example, Chiucchi and Gibbs (2010) contrasted these two estimates to investigate effects of habitat fragmentation on rattlesnake population connectivity and concluded that habitat fragmentation had had minimal effect on genetic connectivity of these populations. Our results show that, if the connectivity had been reduced substantially in recent times, (i) the historical connectivity estimate is likely to underestimate the true historical connectivity before the reduction, and (ii) the contemporary estimate may overestimate actual contemporary connectivity. These two effects, in concert, will result in a relatively small difference between the two estimates, leading to the potentially erroneous conclusion that there is no significant difference in connectivity between two time periods. This could provide an explanation for the paradoxical finding that rates of contemporary and historical gene flow were similar for rattlesnake populations, even though the habitat has been highly fragmented in recent times (Chiucchi & Gibbs 2010). Therefore, we recommend caution, especially when it involves an organism of conservation interest.

2.5.4 Thinking about time since change relative to the effective population size

Another key message for conservation biologists and molecular ecologists is that the time since the change in connectivity should be thought of in terms of effective population size generations (i.e., $N_e$ generations) rather than just in generations. Because the time needed for the fixation of one allele depends on the effective population size (Kimura & Ohta 1969), it requires a longer time (in generations) to reflect recent changes in connectivity for a large population than for a
small population. As a result, the error in the estimate for a given time (e.g. when change occurred 100 generation before present) is higher for a population of 1000 than for a population of 100. This is because, when the change occurred 100 generation ago, time since the change for a population of $N_e = 1000$ is $0.1 N_e$ generations ($100/1000$), while time since the change for a population of $N_e = 100$ is $1 N_e$ generations ($100/100$). Although the time passed in nominal generations is equal in this example (i.e., 100 generations), the change is very recent for the large population (only $0.1 N_e$ generations ago) but longer ago for the small population ($1 N_e$ generations ago). Consequently, it is important to think of time since change in connectivity in terms of the effective population size generations.

2.5.5 Summary

In situations where population genetic connectivity has declined recently from historical levels, our results indicate that estimates of recent genetic connectivity are likely to be overestimated by methods such as BayesAss. These overestimates could have important consequences for populations (or species as a whole) of conservation concern where current genetic diversity is low and the potential role of the rescue effect is being evaluated to determine the probability of long-term persistence. Furthermore, our results indicate, when connectivity has been reduced substantially in the recent past, coalescent-based methods are likely to underestimate historical levels of connectivity. When historical connectivity estimates from coalescent-based methods and recent connectivity estimates using disequilibrium-based assignment methods are compared to make inferences about connectivity over time, the change in connectivity between time periods is likely to be underestimated. Therefore, researchers and managers need to be cautious about interpreting their results, especially when it involves a population or species of conservation concern. Finally, these findings highlight the need for better genetic methods to
estimate contemporary connectivity because recent, reduced levels of connectivity are expected to become increasingly common for many species because of habitat changes.

2.6 Acknowledgements

The authors thank Stephen Wright for valuable feedback on study design, the interpretation of results, and comments on the manuscript; Ben Gilbert for helpful feedback on interpreting our results; and Ken Minns for providing valuable comments on the manuscript. We thank Robert Williamson and Emily Josephs for helping with Python scripts for analysis. This work was made possible by use of Sharcnet High Performance Computer cluster for computation. Samarasin was supported by NSERC grants (to H. Rodd and B. Shuter) and by the Department of Ecology & Evolutionary Biology at the University of Toronto.

2.7 Literature cited


Chapter 2: Genetic connectivity in a changing world


Chapter 2: Genetic connectivity in a changing world


Chapter 2: Genetic connectivity in a changing world


2.8 Table

Table 2.1: The ratio of mean square error when connectivity change was 5 generation ago (i.e., recent) versus when connectivity change was 500 generations ago (i.e., historic).

<table>
<thead>
<tr>
<th>Change in gene flow</th>
<th>Population size</th>
<th>Migrate-n MSE_{5}/MSE_{500}</th>
<th>BayesAss MSE_{5}/MSE_{500}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nm = 1 to 0</td>
<td>100</td>
<td>18.3</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>21.1</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.8</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Nm = 1 to 0.25</td>
<td>100</td>
<td>0.6</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>3.6</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>1.3</td>
<td>1.1</td>
</tr>
<tr>
<td>Nm = 1 to 0.5</td>
<td>100</td>
<td>0.7</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.5</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>1.3</td>
<td>0.8</td>
</tr>
<tr>
<td>Nm = 1 to 0.75</td>
<td>100</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.8</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.1</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>1.0</td>
<td>0.9</td>
</tr>
</tbody>
</table>
2.9 Figures

Figure 2.1: A schematic showing an example of connectivity among four populations of a hypothetical species pre and post road construction. a) Four populations exchange migrants with no barrier to restrict migration among patches (prior to road construction). b) Roads were constructed 50 generations ago and effective migration was reduced to zero from that time until present day.
Figure 2.2: Plot of $F_{st}$ over 5000 generations for a data set with population size of 200 ($N=200$) and a migration rate change from $Nm=1$ to $Nm=0$, 50 generations before the end ($Nm_{before} = 1$, $Nm_{after} = 0$). Note the increase in $F_{st}$ in the last 50 generations as a result of the reduction in migration.
Figure 2.3: The mean square error (panel a) and bias (panel b) of migration rate estimates using BayesAss (dark bar) and Migrate-n (grey bar) when gene flow changed from $N_m=1$ to $N_m=0$ at different points in time (x-axis) with different population sizes ($N=100, 200, 1000, 2000$). Error bars represent the standard error of the mean.
Figure 2.4: Relationship between relative mean square error (RMSE) and magnitude of gene flow change ($\Delta N_m$) for population size of 200 ($N=200$). Error bars represent the standard error of the mean. The figure shows increasing accuracy (i.e., decreasing RMSE) as the magnitude of the gene flow change decrease.
Figure 2.5: Effect of time since the change in connectivity on relative mean square error when time is scaled to the population size used in the simulation. A gene flow change from Nm = 1 to Nm = 0.25 was used to generate the results for this figure.
Chapter 2: Genetic connectivity in a changing world

2.10 Appendix

A 2.1: Mean square error (MSE) and relative mean square error (RMSE) for Migrate-n and BayesAss estimates for all combinations of gene flow change, population size and time since the change in migration.

<table>
<thead>
<tr>
<th>Change in gene flow</th>
<th>Population size (N)</th>
<th>Time since change (generations)</th>
<th>( \text{MSE}_{\text{Migrate}} )</th>
<th>( \text{Bias}_{\text{Migrate}} )</th>
<th>( \text{MSE}_{\text{BayesAss}} )</th>
<th>( \text{Bias}_{\text{BayesAss}} )</th>
<th>( \text{RMSE}_{\text{Migrate}} )</th>
<th>( \text{RBias}_{\text{Migrate}} )</th>
<th>( \text{RMSE}_{\text{BayesAss}} )</th>
<th>( \text{RBias}_{\text{BayesAss}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nm = 1 to 0 [ 1 ]</td>
<td>100 [ 5 ]</td>
<td>[ 0 ]</td>
<td>7.00E-06</td>
<td>2.52E-03</td>
<td>1.48E-05</td>
<td>0.00361806</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>6.82E-06</td>
<td>2.49E-03</td>
<td>1.23E-05</td>
<td>0.00330803</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[ 5 ]</td>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>6.15E-06</td>
<td>2.38E-03</td>
<td>9.55E-06</td>
<td>0.00302269</td>
<td>NA</td>
</tr>
<tr>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>4.17E-06</td>
<td>1.89E-03</td>
<td>8.13E-06</td>
<td>0.0027584</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>[ 500 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>3.46E-06</td>
<td>1.69E-03</td>
<td>7.06E-06</td>
<td>0.00265497</td>
<td>NA</td>
</tr>
<tr>
<td>[ 200 ]</td>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>1.76E-06</td>
<td>1.17E-03</td>
<td>6.95E-06</td>
<td>0.00263456</td>
<td>NA</td>
</tr>
<tr>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>[ 500 ]</td>
<td>[ 200 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>3.82E-07</td>
<td>4.09E-04</td>
<td>7.00E-06</td>
<td>0.00264578</td>
<td>NA</td>
</tr>
<tr>
<td>[ 200 ]</td>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>4.60E-06</td>
<td>2.04E-03</td>
<td>1.11E-05</td>
<td>0.00322466</td>
<td>NA</td>
</tr>
<tr>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>4.85E-06</td>
<td>2.12E-03</td>
<td>1.20E-05</td>
<td>0.00326684</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>4.15E-06</td>
<td>1.92E-03</td>
<td>8.65E-06</td>
<td>0.00289075</td>
<td>NA</td>
</tr>
<tr>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>[ 500 ]</td>
<td>[ 200 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>2.95E-06</td>
<td>1.62E-03</td>
<td>7.42E-06</td>
<td>0.00271578</td>
</tr>
<tr>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>[ 500 ]</td>
<td>[ 200 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>1.82E-06</td>
<td>1.23E-03</td>
<td>7.11E-06</td>
<td>0.00266448</td>
<td>NA</td>
</tr>
<tr>
<td>[ 200 ]</td>
<td>[ 10 ]</td>
<td>[ 20 ]</td>
<td>[ 50 ]</td>
<td>[ 100 ]</td>
<td>[ 200 ]</td>
<td>9.40E-07</td>
<td>8.18E-04</td>
<td>6.95E-06</td>
<td>0.00263541</td>
<td>NA</td>
</tr>
</tbody>
</table>
Chapter 2: Genetic connectivity in a changing world

<table>
<thead>
<tr>
<th>Nm</th>
<th>500</th>
<th>1000</th>
<th>2000</th>
<th>100</th>
<th>200</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.18E-07</td>
<td>2.98E-04</td>
<td>6.98E-06</td>
<td>0.00264021</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>5</td>
<td>2.98E-06</td>
<td>1.70E-03</td>
<td>7.15E-06</td>
<td>0.00266891</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>10</td>
<td>2.96E-06</td>
<td>1.69E-03</td>
<td>7.03E-06</td>
<td>0.00265003</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>20</td>
<td>2.97E-06</td>
<td>1.69E-03</td>
<td>7.02E-06</td>
<td>0.00264838</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>50</td>
<td>2.83E-06</td>
<td>1.65E-03</td>
<td>7.01E-06</td>
<td>0.00264655</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>100</td>
<td>2.60E-06</td>
<td>1.57E-03</td>
<td>7.02E-06</td>
<td>0.00264895</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>200</td>
<td>2.22E-06</td>
<td>1.44E-03</td>
<td>7.03E-06</td>
<td>0.0026506</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>500</td>
<td>1.68E-06</td>
<td>1.18E-03</td>
<td>7.00E-06</td>
<td>0.00264495</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>2000</td>
<td>2.72E-06</td>
<td>1.61E-03</td>
<td>7.00E-06</td>
<td>0.00264526</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>10</td>
<td>2.78E-06</td>
<td>1.63E-03</td>
<td>6.93E-06</td>
<td>0.0026322</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>20</td>
<td>2.72E-06</td>
<td>1.62E-03</td>
<td>7.18E-06</td>
<td>0.00266123</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>50</td>
<td>2.61E-06</td>
<td>1.56E-03</td>
<td>7.00E-06</td>
<td>0.00264577</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>100</td>
<td>2.57E-06</td>
<td>1.56E-03</td>
<td>7.02E-06</td>
<td>0.00264805</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>200</td>
<td>2.31E-06</td>
<td>1.46E-03</td>
<td>6.98E-06</td>
<td>0.00264219</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>500</td>
<td>1.84E-06</td>
<td>1.25E-03</td>
<td>7.00E-06</td>
<td>0.0026444</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Nm = 1 to 0.25

| 100 | 5  | 6.88E-07 | 0.000145 | 1.30E-05 | 0.00209809 | 0.1100168 | 0.05808333 | 2.083569 | 0.8392361 |
| 10  | 5.12E-07 | 0.000126 | 1.02E-05 | 0.00177636 | 0.08186543 | 0.05025 | 1.637341 | 0.7105442 |
| 20  | 5.61E-07 | -0.000144 | 1.11E-05 | 0.00169132 | 0.08983448 | -0.0574917 | 1.779517 | 0.6765268 |
| 50  | 5.67E-07 | -0.00027 | 8.93E-06 | 0.00144046 | 0.09066478 | -0.108025 | 1.429306 | 0.5761838 |
Chapter 2: Genetic connectivity in a changing world

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6.73E-07</td>
<td>-0.000523</td>
<td>5.83E-06</td>
<td>0.00103038</td>
<td>0.1077566</td>
<td>-0.20935</td>
<td>0.9320081</td>
</tr>
<tr>
<td></td>
<td>8.59E-07</td>
<td>-0.000546</td>
<td>6.26E-06</td>
<td>0.00108375</td>
<td>0.1374117</td>
<td>-0.218325</td>
<td>1.001534</td>
</tr>
<tr>
<td></td>
<td>1.16E-06</td>
<td>-0.000732</td>
<td>6.34E-06</td>
<td>0.00083713</td>
<td>0.1856759</td>
<td>-0.29285</td>
<td>1.014536</td>
</tr>
<tr>
<td>200</td>
<td>1.21E-06</td>
<td>0.000903</td>
<td>1.11E-05</td>
<td>0.00250112</td>
<td>0.774933</td>
<td>0.722633</td>
<td>7.097222</td>
</tr>
<tr>
<td></td>
<td>8.33E-07</td>
<td>0.000721</td>
<td>8.55E-06</td>
<td>0.00227412</td>
<td>0.5433275</td>
<td>0.774793</td>
<td>1.014536</td>
</tr>
<tr>
<td></td>
<td>3.79E-07</td>
<td>0.000128</td>
<td>7.64E-06</td>
<td>0.00201881</td>
<td>0.2424634</td>
<td>0.533147</td>
<td>1.014536</td>
</tr>
<tr>
<td></td>
<td>3.39E-07</td>
<td>-6.96E-05</td>
<td>6.69E-06</td>
<td>0.00207322</td>
<td>0.2171215</td>
<td>-0.0557</td>
<td>4.28342</td>
</tr>
<tr>
<td>1000</td>
<td>2.19E-06</td>
<td>0.001449</td>
<td>7.95E-06</td>
<td>0.00265352</td>
<td>35.10997</td>
<td>5.794917</td>
<td>127.2551</td>
</tr>
<tr>
<td></td>
<td>2.17E-06</td>
<td>0.001448</td>
<td>7.11E-06</td>
<td>0.00256525</td>
<td>34.79384</td>
<td>5.790333</td>
<td>113.7129</td>
</tr>
<tr>
<td></td>
<td>2.15E-06</td>
<td>0.001433</td>
<td>7.18E-06</td>
<td>0.00253298</td>
<td>34.42793</td>
<td>5.730333</td>
<td>114.8444</td>
</tr>
<tr>
<td></td>
<td>2.04E-06</td>
<td>0.001397</td>
<td>6.66E-06</td>
<td>0.00249148</td>
<td>32.70403</td>
<td>5.587</td>
<td>106.5398</td>
</tr>
<tr>
<td>2000</td>
<td>1.99E-06</td>
<td>0.001365</td>
<td>8.31E-06</td>
<td>0.00262571</td>
<td>31.86088</td>
<td>5.458167</td>
<td>132.982</td>
</tr>
<tr>
<td></td>
<td>1.71E-06</td>
<td>0.001265</td>
<td>6.97E-06</td>
<td>0.00252525</td>
<td>27.41676</td>
<td>5.059167</td>
<td>111.5225</td>
</tr>
<tr>
<td></td>
<td>1.46E-06</td>
<td>0.001154</td>
<td>7.07E-06</td>
<td>0.00249569</td>
<td>23.34745</td>
<td>4.616417</td>
<td>113.1285</td>
</tr>
<tr>
<td>5000</td>
<td>2.38E-06</td>
<td>0.001504</td>
<td>7.30E-06</td>
<td>0.00261641</td>
<td>152.0509</td>
<td>12.03083</td>
<td>467.1371</td>
</tr>
<tr>
<td></td>
<td>2.41E-06</td>
<td>0.001505</td>
<td>7.15E-06</td>
<td>0.0026179</td>
<td>154.154</td>
<td>12.03683</td>
<td>457.5252</td>
</tr>
</tbody>
</table>
Chapter 2: Genetic connectivity in a changing world

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2.28E-06</td>
<td>0.001477</td>
<td>7.00E-06</td>
<td>0.00258689</td>
<td>146.1632</td>
<td>11.813</td>
</tr>
<tr>
<td>50</td>
<td>2.25E-06</td>
<td>0.001466</td>
<td>7.34E-06</td>
<td>0.00261874</td>
<td>143.9201</td>
<td>11.728</td>
</tr>
<tr>
<td>100</td>
<td>2.21E-06</td>
<td>0.001437</td>
<td>6.35E-06</td>
<td>0.00251906</td>
<td>141.5742</td>
<td>11.4935</td>
</tr>
<tr>
<td>200</td>
<td>2.09E-06</td>
<td>0.001393</td>
<td>7.44E-06</td>
<td>0.00263808</td>
<td>133.5073</td>
<td>11.1425</td>
</tr>
<tr>
<td>500</td>
<td>1.86E-06</td>
<td>0.001284</td>
<td>6.35E-06</td>
<td>0.00251921</td>
<td>118.8804</td>
<td>10.27267</td>
</tr>
</tbody>
</table>

Nm = 1 to 0.50

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>5</td>
<td>5.96E-06</td>
<td>-0.002309</td>
<td>1.93E-05</td>
<td>0.00077379</td>
<td>0.2385616</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6.39E-06</td>
<td>-0.002432</td>
<td>1.75E-05</td>
<td>0.0006425</td>
<td>0.2554381</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>6.74E-06</td>
<td>-0.002506</td>
<td>1.97E-05</td>
<td>0.00038398</td>
<td>0.2694078</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.78E-06</td>
<td>-0.002492</td>
<td>1.48E-05</td>
<td>7.69E-05</td>
<td>0.2710472</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>7.24E-06</td>
<td>-0.002597</td>
<td>9.16E-06</td>
<td>-0.0006743</td>
<td>0.2894882</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>7.77E-06</td>
<td>-0.00267</td>
<td>1.50E-05</td>
<td>-0.0006388</td>
<td>0.3108437</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>8.35E-06</td>
<td>-0.002786</td>
<td>1.44E-05</td>
<td>-9.77E-05</td>
<td>0.3338551</td>
</tr>
</tbody>
</table>

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>5</td>
<td>6.17E-07</td>
<td>-0.000424</td>
<td>1.20E-05</td>
<td>0.00185291</td>
<td>0.09879468</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>5.41E-07</td>
<td>-0.000406</td>
<td>9.59E-06</td>
<td>0.0016421</td>
<td>0.0864884</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>5.52E-07</td>
<td>-0.00043</td>
<td>1.37E-05</td>
<td>0.00175663</td>
<td>0.0882775</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>6.98E-07</td>
<td>-0.000615</td>
<td>9.36E-06</td>
<td>0.00132734</td>
<td>0.1117008</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>7.92E-07</td>
<td>-0.00067</td>
<td>5.80E-06</td>
<td>0.00131215</td>
<td>0.126748</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>8.53E-07</td>
<td>-0.000704</td>
<td>9.97E-06</td>
<td>0.00134898</td>
<td>0.136448</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>1.16E-06</td>
<td>-0.000874</td>
<td>6.41E-06</td>
<td>0.00117538</td>
<td>0.1861461</td>
</tr>
</tbody>
</table>
Chapter 2: Genetic connectivity in a changing world

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>1.56E-06</td>
<td>0.001207</td>
<td>6.17E-06</td>
<td>0.00233412</td>
<td>6.255563</td>
</tr>
<tr>
<td>10</td>
<td>1.51E-06</td>
<td>0.00118</td>
<td>5.44E-06</td>
<td>0.00225487</td>
<td>6.027159</td>
<td>2.360917</td>
</tr>
<tr>
<td>20</td>
<td>1.50E-06</td>
<td>0.00119</td>
<td>6.91E-06</td>
<td>0.00240532</td>
<td>5.988886</td>
<td>2.380292</td>
</tr>
<tr>
<td>50</td>
<td>1.37E-06</td>
<td>0.001139</td>
<td>7.30E-06</td>
<td>0.00245819</td>
<td>5.498675</td>
<td>2.277708</td>
</tr>
<tr>
<td>100</td>
<td>1.42E-06</td>
<td>0.00115</td>
<td>6.16E-06</td>
<td>0.00232947</td>
<td>5.699816</td>
<td>2.198033</td>
</tr>
<tr>
<td>200</td>
<td>1.30E-06</td>
<td>0.001095</td>
<td>7.67E-06</td>
<td>0.00247642</td>
<td>5.988886</td>
<td>2.380292</td>
</tr>
<tr>
<td>500</td>
<td>1.24E-06</td>
<td>0.001062</td>
<td>5.90E-06</td>
<td>0.00229735</td>
<td>5.498675</td>
<td>2.277708</td>
</tr>
<tr>
<td>2000</td>
<td>5</td>
<td>2.07E-06</td>
<td>0.001402</td>
<td>5.99E-06</td>
<td>0.00242956</td>
<td>33.16353</td>
</tr>
<tr>
<td>10</td>
<td>1.96E-06</td>
<td>0.001357</td>
<td>7.09E-06</td>
<td>0.00254349</td>
<td>31.29493</td>
<td>5.426667</td>
</tr>
<tr>
<td>20</td>
<td>1.95E-06</td>
<td>0.001354</td>
<td>6.11E-06</td>
<td>0.00243928</td>
<td>31.29493</td>
<td>5.426667</td>
</tr>
<tr>
<td>50</td>
<td>1.92E-06</td>
<td>0.001344</td>
<td>6.04E-06</td>
<td>0.00242583</td>
<td>30.79063</td>
<td>5.3755</td>
</tr>
<tr>
<td>100</td>
<td>1.99E-06</td>
<td>0.001365</td>
<td>5.76E-06</td>
<td>0.00239988</td>
<td>31.90979</td>
<td>5.45925</td>
</tr>
<tr>
<td>200</td>
<td>1.78E-06</td>
<td>0.001281</td>
<td>6.66E-06</td>
<td>0.00249568</td>
<td>28.48495</td>
<td>5.123917</td>
</tr>
<tr>
<td>500</td>
<td>1.55E-06</td>
<td>0.001172</td>
<td>7.19E-06</td>
<td>0.00255315</td>
<td>24.77449</td>
<td>4.689917</td>
</tr>
</tbody>
</table>

Nm = 1 to 0.75

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>5</td>
<td>2.42E-05</td>
<td>-0.004854</td>
<td>2.72E-05</td>
<td>-0.00088</td>
<td>0.4294447</td>
</tr>
<tr>
<td>10</td>
<td>2.39E-05</td>
<td>-0.004836</td>
<td>2.49E-05</td>
<td>-0.0011242</td>
<td>0.4244292</td>
<td>-0.6448278</td>
</tr>
<tr>
<td>20</td>
<td>2.55E-05</td>
<td>-0.005014</td>
<td>3.04E-05</td>
<td>-0.0013398</td>
<td>0.4538605</td>
<td>-0.6685361</td>
</tr>
<tr>
<td>50</td>
<td>2.49E-05</td>
<td>-0.004927</td>
<td>2.50E-05</td>
<td>-0.0014758</td>
<td>0.4433696</td>
<td>-0.6568694</td>
</tr>
<tr>
<td>100</td>
<td>2.41E-05</td>
<td>-0.004851</td>
<td>3.10E-05</td>
<td>-0.0010808</td>
<td>0.4278738</td>
<td>-0.6467556</td>
</tr>
</tbody>
</table>
### Chapter 2: Genetic connectivity in a changing world

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>2.43E-05</td>
<td>-0.004861</td>
<td>2.29E-05</td>
<td>-0.0015291</td>
<td>0.4312128</td>
<td>-0.6481083</td>
<td>0.4073447</td>
<td>-0.2038792</td>
</tr>
<tr>
<td>500</td>
<td>2.51E-05</td>
<td>-0.004959</td>
<td>2.00E-05</td>
<td>-0.00212</td>
<td>0.4456086</td>
<td>-0.6612573</td>
<td>0.3557563</td>
<td>-0.2826604</td>
</tr>
<tr>
<td>200</td>
<td>2.82E-06</td>
<td>-0.001555</td>
<td>1.81E-05</td>
<td>0.00159086</td>
<td>0.2003738</td>
<td>-0.4145722</td>
<td>1.284217</td>
<td>0.4242302</td>
</tr>
<tr>
<td></td>
<td>-0.001606</td>
<td>2.92E-06</td>
<td>1.19E-05</td>
<td>0.00097705</td>
<td>0.2073227</td>
<td>-0.4282611</td>
<td>0.8430645</td>
<td>0.2605456</td>
</tr>
<tr>
<td>20</td>
<td>2.98E-06</td>
<td>-0.001614</td>
<td>1.25E-05</td>
<td>0.00104876</td>
<td>0.2119061</td>
<td>-0.4302833</td>
<td>0.8879217</td>
<td>0.2796687</td>
</tr>
<tr>
<td>50</td>
<td>3.15E-06</td>
<td>-0.001665</td>
<td>1.43E-05</td>
<td>0.00133755</td>
<td>0.2239977</td>
<td>-0.4441056</td>
<td>1.014269</td>
<td>0.3566805</td>
</tr>
<tr>
<td>100</td>
<td>3.42E-06</td>
<td>-0.001762</td>
<td>1.15E-05</td>
<td>0.00079929</td>
<td>0.243127</td>
<td>-0.46975</td>
<td>0.8151706</td>
<td>0.2131427</td>
</tr>
<tr>
<td>200</td>
<td>3.51E-06</td>
<td>-0.001786</td>
<td>5.56E-06</td>
<td>0.00031939</td>
<td>0.2493415</td>
<td>-0.4761333</td>
<td>0.3956151</td>
<td>0.0851701</td>
</tr>
<tr>
<td>500</td>
<td>3.32E-06</td>
<td>-0.001727</td>
<td>1.33E-05</td>
<td>0.00059705</td>
<td>0.2362873</td>
<td>-0.4605056</td>
<td>0.9464352</td>
<td>0.1592134</td>
</tr>
<tr>
<td>1000</td>
<td>5</td>
<td>1.02E-06</td>
<td>0.000969</td>
<td>8.90E-06</td>
<td>0.00243578</td>
<td>1.815674</td>
<td>1.291833</td>
<td>15.82253</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.06E-06</td>
<td>0.000984</td>
<td>4.99E-06</td>
<td>0.00207325</td>
<td>1.892787</td>
<td>1.311944</td>
<td>8.879319</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1.05E-06</td>
<td>0.000974</td>
<td>6.31E-06</td>
<td>0.0022296</td>
<td>1.858801</td>
<td>1.298639</td>
<td>11.22168</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>1.00E-06</td>
<td>0.000952</td>
<td>6.66E-06</td>
<td>0.00222054</td>
<td>1.7815</td>
<td>1.269778</td>
<td>11.84385</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>9.92E-07</td>
<td>0.000938</td>
<td>6.18E-06</td>
<td>0.00220517</td>
<td>1.763906</td>
<td>1.25025</td>
<td>10.98076</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>9.16E-07</td>
<td>0.000895</td>
<td>5.84E-06</td>
<td>0.00217206</td>
<td>1.628867</td>
<td>1.193528</td>
<td>10.39004</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>9.34E-07</td>
<td>0.000911</td>
<td>6.35E-06</td>
<td>0.00222584</td>
<td>1.661187</td>
<td>1.21525</td>
<td>11.28558</td>
</tr>
<tr>
<td></td>
<td>2000</td>
<td>5</td>
<td>1.62E-06</td>
<td>0.001226</td>
<td>5.72E-06</td>
<td>0.00233219</td>
<td>11.53034</td>
<td>3.269167</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1.68E-06</td>
<td>0.001245</td>
<td>6.44E-06</td>
<td>0.00241235</td>
<td>11.92246</td>
<td>3.320944</td>
<td>45.77745</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1.65E-06</td>
<td>0.001234</td>
<td>6.35E-06</td>
<td>0.00240453</td>
<td>11.74784</td>
<td>3.290556</td>
<td>45.19064</td>
</tr>
</tbody>
</table>
Chapter 2: Genetic connectivity in a changing world

<table>
<thead>
<tr>
<th></th>
<th>1.57E-06</th>
<th>0.001204</th>
<th>6.97E-06</th>
<th>0.00247337</th>
<th>11.15552</th>
<th>3.209944</th>
<th>49.59066</th>
<th>6.595639</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1.65E-06</td>
<td>0.00123</td>
<td>6.04E-06</td>
<td>0.00236934</td>
<td>11.69794</td>
<td>3.280667</td>
<td>42.92496</td>
<td>6.318234</td>
</tr>
<tr>
<td>100</td>
<td>1.58E-06</td>
<td>0.001199</td>
<td>6.46E-06</td>
<td>0.00241958</td>
<td>11.21428</td>
<td>3.197</td>
<td>45.9205</td>
<td>6.4522</td>
</tr>
<tr>
<td>200</td>
<td>1.56E-06</td>
<td>0.001188</td>
<td>6.06E-06</td>
<td>0.00236478</td>
<td>11.07304</td>
<td>3.166889</td>
<td>43.06195</td>
<td>6.306073</td>
</tr>
</tbody>
</table>
Chapter 3

After 100 years: Hydroelectric dam-induced life-history divergence and population genetic changes in Sockeye salmon

3.1 Abstract

Understanding the impact of barriers and habitat fragmentation on the ecology and genetics of species is of broad interest to many biologists. In aquatic systems, hydroelectric dams often present an impenetrable barrier to migratory fishes and can have negative effects on their persistence. Hydroelectric dams constructed in the Coquitlam and Alouette rivers in the Fraser River drainage (British Columbia, Canada) in the early 1900’s were thought to have led to the loss of anadromous Sockeye salmon from both rivers by the 1930’s. Recent water release programs for both reservoirs resulted in the unexpected downstream migration of juvenile salmon and the subsequent upstream migration of adults towards the reservoir two years later. Investigation of the origins of the upstream-migrating adults by Godbout et al. (2011) suggested that returning adults originated from these reservoirs and not from other nearby populations. Here we investigate the genetic distinction between migratory “sea-run” individuals and presumed “resident” individuals within the Alouette and Coquitlam reservoirs. We also compare historical and contemporary genetic connectivity among eleven Lower Fraser River Sockeye sites to identify recent changes that might have been influenced by anthropogenic activities. Our molecular genetic analyses show a genetic distinction between the sea-run and resident individuals from the Coquitlam reservoir and population splitting time estimates suggest a very recent divergence. Gene flow comparisons suggest a general decline in gene flow among most sites, with a few interesting exceptions. In summary, our results suggest (i) early-stage
Chapter 3: Hydroelectric dam induced life-history divergence
divergence between life history forms of Sockeye salmon within the Coquitlam reservoir, and (ii) recent changes in genetic connectivity among Lower Fraser River populations, both with potentially important implications for the management of salmon populations.

3.2 Introduction
The genetic structure of populations can be influenced by both biological characteristics of the organism (e.g., dispersal behaviour, mating system) and characteristics of the environment (e.g., landscape features, predator distribution). Landscape features such as barriers (e.g., dams) that restrict gene flow can reduce genetic diversity (Wofford et al. 2005) and may cause populations that would otherwise be panmictic to differentiate. When populations become fragmented, effective population size is usually reduced and, because of this and reduced immigration, the probability of local extirpation is increased (Hanski & Gilpin 1996; Lande 1998). Dam-induced fragmentation of freshwater habitats and its negative consequences have been documented in many species of fish. Dams have been linked to increased juvenile mortality (Budy et al. 2002; Williams et al. 2001), reduced genetic diversity because of reduced gene flow, and increased genetic drift because of reduced population size (e.g., (Horreo et al. 2011; Wofford et al. 2005). Dams have led to other negative effects such as a reduction in biomass of white spotted charr (*Salvelinus leucomaenis*), which occurred when the anadromous form converted to the resident form because of a dam (Morita et al. 2009). Dams can have indirect costs, including physiological ones; for example, rainbow trout (*Oncorhynchus mykiss*) in dammed sites continue to undergo smoltification (a physiological change in preparation for the marine environment), likely at a significant energetic cost, even though there is no benefit in doing so because of the dam (Holecek et al. 2012).
Chapter 3: Hydroelectric dam induced life-history divergence

The Fraser River drainage in British Columbia, Canada is home to millions of Sockeye salmon, distributed across about 90 populations (Withler et al. 2000). Here we present an investigation of the effects of hydroelectric dams on Sockeye salmon in the Lower Fraser River. Sockeye salmon is typically anadromous; it is born in fresh water and migrates to the Pacific Ocean where it matures and returns to its freshwater natal site to spawn. There is also a non-anadromous form of Sockeye salmon (called kokanee) that spends its entire life in fresh water (typically lakes) without migrating to the ocean. Analyses by Ricker (1940) and Foote et al. (1989) suggested that the kokanee form evolved from the anadromous form, likely as a consequence of impenetrable barriers that prevented ocean migration (Wood & Foote 1996). As juveniles, anadromous and non-anadromous Sockeye salmon are morphologically indistinguishable from one another. As adults, there are differences in body size (Wood 1995) and the number of gill rakers (Nelson 1968).

Unfortunately, Sockeye salmon runs have been declining since the 1990s (Cooke et al. 2004) and there have been very serious concerns about the status of Canadian Sockeye salmon populations in recent years. Multiple causes have been implicated in the decline in Sockeye salmon including habitat loss and degradation, overfishing, disease and parasites from fish farms, and climate change (Cooke et al. 2004; Healey 2011; Rand et al. 2012; Slaney et al. 1996). Barriers that prevent Sockeye salmon from reaching their natal spawning site, including dams, have also contributed to declines in some Sockeye salmon populations, and at least two populations of the migrating, anadromous Sockeye salmon have been extirpated due to dams (Slaney et al. 1996). Indirectly, barriers such as dams may influence the water flow regime, changing the water chemistry and the chemical cues that salmon use to find their natal site. Therefore, direct and
Chapter 3: Hydroelectric dam induced life-history divergence

indirect effect of barriers that cause individuals to reach a different site than the natal site (i.e., stray) can have severe negative fitness consequences because Sockeye salmon have strong local adaptations (Taylor 1991; Wood 1995).

In the Lower Fraser River, construction of hydroelectric dams were completed in the Coquitlam and Alouette rivers by 1914 and 1928, respectively, which made spawning sites in the reservoirs up-river from dams (i.e., previously spawning sites in the lake) inaccessible to anadromous fish. Historical records indicate that the anadromous Sockeye salmon disappeared from both rivers by 1930’s (Hirst 1991; Slaney et al. 1996), and that the kokanee form has been residing in both reservoirs above the dam since at least the 1950’s. Historical records do not indicate the existence of kokanee in either lake prior to the construction of these dams, and neither lake was stocked with Sockeye salmon after dams were built (Godbout et al. 2011). Interestingly, during recent experimental water releases from these reservoirs (started in 2005), some juvenile salmon were observed migrating downstream towards the ocean. Two years later, for the first time since the 1930’s, adult Sockeye were observed swimming upstream in both the Alouette and Coquitlam rivers towards the dams. Godbout et al. (2011) investigated the origins of these upstream migrating adults using genetic and stable isotope analyses to determine whether these individuals originated from their respective reservoir (i.e. from the presumed kokanee) or were from a nearby population (i.e., strays). Their results indicated that upstream migrating adults originated from the respective reservoirs and that these migrants were likely the descendants of anadromous Sockeye salmon that were trapped by the dam construction (Godbout et al. 2011). This interesting finding, along with the observation that some salmon migrated downstream but others remained in the reservoir, raises several general questions about the evolution and flexibility of life history types in Sockeye salmon. For example, is migratory versus resident
Chapter 3: Hydroelectric dam induced life-history divergence

behaviour phenotypically plastic or is there a genetic basis for some component of the differences between the anadromous and kokanee forms? Are there different gene pools of Sockeye salmon within the Alouette and Coquitlam reservoirs? If so, what is the origin of these two life history types? Are they all descendants of anadromous Sockeye salmon that were trapped after dam construction as believed? Or did kokanee exist in these two lakes prior to dam construction such that each reservoir has two types of individuals: those descended from trapped anadromous Sockeye salmon, and historical kokanee that may have existed before the dam construction?

In this study, we investigate whether there is any genetic distinction between the juveniles that migrated downstream towards the ocean (called sea-run individuals) and individuals that remained in the reservoir (called resident individuals) for both the Alouette and Coquitlam reservoirs. We also investigate possible indirect effects of dams on the natal migration of Sockeye salmon in the Lower Fraser River by investigating the historical and contemporary genetic connectivity among sites. By comparing long-term, historical gene flow estimates against contemporary gene flow estimates, we expect to ascertain changes in population genetic connectivity between historical and current times that can be attributed mainly to dams (the main change in recent past) or to other unidentified causes in the Lower Fraser River region.

3.3 Materials and Methods

3.3.1 DNA sampling
Approximately 40-90 individuals were sampled from each of 11 sites in the lower Fraser River region (shown in Figure 1). The term “Sockeye” is conventionally used to refer to the anadromous type and the “kokanee” is used for resident type. The 11 sites were: Alouette
Chapter 3: Hydroelectric dam induced life-history divergence

reservoir “sea-run” individuals; Alouette reservoir “resident” individuals; Coquitlam reservoir “sea-run” individuals; Coquitlam reservoir “resident” individuals; Big Silver Creek Sockeye; Birkenhead River Sockeye; Pitt River Sockeye; Harrison Rapids Sockeye; Weaver Creek Sockeye; Widgeon Slough Sockeye; and, Stave Lake kokanee. “Sea-run” individuals were juveniles that migrated downstream towards the ocean during experimental water release from the Alouette or Coquitlam reservoir. These individuals were caught with rotary screw traps in the river, a few kilometers downstream of the dams and reservoirs. “Residents” were sampled from the interior of the reservoir and are presumed to be the “resident” type. Please note that individuals in the reservoir can only travel downstream during water release (i.e. human-assisted) and any individuals downstream of the dam cannot enter the reservoir because the dam does not have a fish ladder. All individuals were genotyped at 14 previously characterized microsatellite loci: Ots2, Ots3 (Banks et al. 1999); Ots103 (Nelson & Beacham 1999); Oki1a, Oki1b, Oki6, Oki10, Oki16, Oki29 (Smith et al. 1998); 1b, 3dre, i1, One8 (Scribner et al. 1996), and, Omy77 (Morris et al. 1996).

3.3.2 Measures of genetic diversity and differentiation

To determine whether the genotype data conformed to Hardy-Weinberg and linkage equilibrium assumptions, deviations from Hardy-Weinberg and linkage equilibrium were evaluated using Arlequin 3.5.1.3 ((Excoffier & Lischer 2010). Allelic richness, observed (H_o) and expected heterozygosity (H_e) was calculated using Arlequin. Genetic differentiation between populations was assessed by calculating the pairwise Fst over all loci and the deviation from zero was evaluated using 10000 permutations tests with Bonferroni correction. We investigated isolation-by-distance using the Mantel test implemented in GenePop 4.1.4 (Rousset 2008). The differentiation between sea-run and resident individuals in the Alouette and Coquitlam reservoirs
Chapter 3: Hydroelectric dam induced life-history divergence

was assessed using the genic differentiation test implemented in GenePop 4.1.4 (Rousset 2008). The riverine distances between sites were used in the analysis because that would be the distance salmon would need to travel if an individual were to migrate to a different site. We estimated effective population size \( (N_e) \) using the SibShip method (Wang 2009), implemented in the program Colony 2.0.5.0. (Wang 2009; Wang 2004). The Sibship method estimates effective population size by using frequencies of sib groups in a sample, inferred from multilocus genotypes (Sibship assignment). \( N_e \) was estimated assuming polygamy (Hauser et al. 2011; Mehranvar 2002), allowing for inbreeding, and using the full likelihood method with the complexity prior option. Five replicate runs were conducted for each site, starting with different seed numbers, to assess convergence and consistency in results. The reported \( N_e \) estimate is the average of five runs.

### 3.3.3 Genetic bottlenecks

One possible explanation for a change in gene flow \( (N_e m) \) is that effective population \( (N_e) \) has changed over time, rather than a change in the migration rate \( (m) \). Therefore, we looked for evidence of genetic bottlenecks using BOTTLENECK v.1.2.02 (Piry et al. 1999), which detects a recent reduction in effective population size by testing for excess heterozygosity within the population; a reduction in effective population size causes allelic number to decrease more than the heterozygosity, resulting in a higher level of heterozygosity than expected based on mutation-drift equilibrium. We assumed a two-phase mutation model with 90% stepwise mutations and 10% multiple-step mutations; the variance among multiple steps was set to 12 as recommended for microsatellite data by Piry et al. (1999). We also analyzed the data assuming a two-phase mutation model with 80% and 95% stepwise mutations to determine the robustness of the results.
Significance of heterozygosity excess over all loci was determined with a one-tailed Wilcoxon
sign rank test.

### 3.3.4 Population genetic structure

We used two methodologically different approaches for inferring genetic clusters (i.e., distinct populations) in the Lower Fraser River Sockeye salmon sites. First, the program STRUCTURE 2.3.4 (Falush *et al.* 2003; Pritchard *et al.* 2000) infers the number of genetic clusters in a collection of genotypes using the Bayesian Markov chain Monte Carlo (MCMC) approach to maximize Hardy-Weinberg and linkage equilibrium. Second, the discriminant analysis of principal components (DAPC) method (Jombart *et al.* 2010) uses a multivariate approach; genotype data are transformed into principal components and then discriminant analysis is used to maximize among-population variation to graphically represent genetic clusters.

Genotype data from all 11 sites were included in the first analysis using STRUCTURE. We tested from K=1 to K=11 predefined populations and performed 10 replicate runs at each K value, running for 500,000 iterations after discarding the first 150,000 iterations as burn-in. We used correlated allele frequencies under the admixture ancestry model. The most likely number of clusters, given the data, was determined using the ln Pr (X/K) and the ΔK approach (Evanno *et al.* 2005) implemented in Structure Harvester v0.56.3 (Earl & vonHoldt 2012). To deal with the multimodality of utilizing multiple independent runs, we used CLUMPP v.1.1.2 (Jakobsson & Rosenberg 2007) to permute the admixture coefficients for the runs with the chosen K-value using the “Greedy” algorithm with 1,000 random input orders. DISTUCT v.1.1 (Rosenberg 2004) was used to visualize the output from CLUMPP. We investigated fine-scale genetic structure for “sea-run” and “resident” individuals from the Alouette and Coquitlam reservoirs in
Chapter 3: Hydroelectric dam induced life-history divergence

separate analyses. We tested from K=1 to K=3, and all other analysis parameters used in STRUCTURE were the same as those described above.

The genetic clustering pattern of Lower Fraser River individuals was also investigated using the discriminant analysis of principal components (DAPC) method implemented in the R package ADEGENET 1.3-6 (Jombart et al. 2010). We retained the first 150 principal components (which contained ~95% of variation) and ran 1,000,000 iterations of the K-means algorithm to assess the optimal number of genetic clusters (K) for the data using Bayesian Information Criterion (BIC). The first four discriminant functions were retained for analysis for DAPC analysis. In a separate analysis, the fine-scale genetic structure of fish from the Alouette and Coquitlam reservoirs were investigated. The first 80 principal components were retained for running the K-means algorithm to assess the optimal number of genetic clusters (K) using Bayesian Information Criterion (BIC). The first two discriminant functions were retained for DAPC analysis.

3.3.5 Genetic connectivity

To assess genetic connectivity among the 11 sites, we used two different approaches. Historical, long-term genetic connectivity among sites was compared with contemporary (recent) genetic connectivity by using the coalescent-based gene flow estimator Migrate-n (Beerli 2006) and disequilibrium-based BayesAss program (Wilson & Rannala 2003) respectively. Migrate-n uses a coalescent framework with Markov Chain Monte Carlo (MCMC) sampling to estimate historical migration rates scaled by mutation rate \( M = m/\mu \) and mutation-scaled effective population size \( \Theta = 4N_e\mu \). All 11 sites were used to estimate gene flow (full migration model) among them. Analyses were performed using the following parameters in the program:
Brownian motion with the relative mutation rate option for microsatellite data, sampling 2000,000 steps at 100 step increments and the first 20,000 genealogies discarded as burn-in. We used the UPGMA tree as the starting genealogy and used static heating with six temperatures as specified by the # command in the program for even spacing (temperatures were 1.00, 1.25, 1.67, 2.5, 5.00, 1000000.0). After multiple independent runs, we used results from the initial analyses as starting values for theta and migration rates to increase the accuracy of the parameter estimates further. Posterior distribution, autocorrelation, and effective sample size (ESS) were checked to judge convergence.

We estimated the contemporary migration rate between the sites using the Bayesian MCMC method implemented in BayesAss 1.3 (Wilson & Rannala 2003). This method does not assume that populations are in Hardy-Weinberg equilibrium and incorporates separate inbreeding coefficients for each population. We performed $50 \times 10^6$ iterations with a sampling frequency of 2000 and discarded the first $10 \times 10^6$ iterations as burn-in. The delta values were initially set to 0.15 and that yielded accepted proposed changes in parameters between 40% and 60% of the total chain length, as recommended by authors. We performed 10 independent runs with different seed numbers and the run with the minimum Bayesian deviance was used for reporting the migration rate estimates (see Faubet et al. 2007 and Spiegelhalter et al. 2002 for more detail). To test the robustness of the results, we ran analyses with delta values of 0.1 and 0.2; they provided similar estimates.
3.3.6 Origins of sea-run and resident types

Because the DAPC analysis suggested a distinction between “sea-run” and “resident” individuals, we next assessed the origins of these two types by estimating population splitting time ($t$) using the Isolation with Migration (IM) model (Hey & Nielsen 2004; Nielsen & Wakeley 2001) implemented in the IMa2 program (Hey 2010). If both sea-run and resident individuals from the reservoir are descendants of the original, trapped ancestral Sockeye salmon, population splitting time would be less than 100 years (i.e., after dam construction). However, if these two forms originated from separate Sockeye lineages that existed in the original lake prior to dam construction, population splitting time would be greater than 100 years (i.e., separate origins traced back before dam construction). We applied a stepwise mutation model (SMM) for microsatellite data and the locus specific mutation rate was estimated by dividing the locus specific theta ($\theta = 4N\mu$) by four times the effective population size estimated using the SibShip method (Wang 2009) ($\mu = \theta / 4N$). This assumes that historical effective population sizes are not substantially different from effective population sizes estimated by the SibShip method. The mutation rate (per generation) was converted to a value ‘per year’ by assuming a generation time of 4 years (Reed et al. 2011). Many trial runs (>20) were conducted to explore the parameter space in order to obtain good upper bounds for parameters of interest and to determine search parameters that would lead to good MCMC mixing, determined by high swap rates and no clear directional trends in likelihood or population splitting time parameter over time. The final analysis was conducted with 350 chains with geometric heating (ha = 0.999, hb= 0.3), a burn-in of 1,500,000 steps and 10,000 genealogies in 100 step intervals saved following burn-in. We completed five independent runs with different seed numbers and combined the results in L-mode.
3.4 Results

3.4.1 Genetic diversity and differentiation
In general, within-site genetic diversity was moderately high as indicated by both the levels of heterozygosity and mean numbers of alleles per locus (Table 1). Observed heterozygosity ranged from a low of 0.52 for the Widgeon Slough site to a high of 0.72 for the Harrison site; mean number of alleles was lowest for the Widgeon Slough site (k = 4.29) and highest for the Pitt site (k = 10.57). None of the population-locus combinations showed significant deviation from Hardy-Weinberg equilibrium or linkage equilibrium after controlling for family-wise error rate using Bonferroni correction. The observed heterozygosity levels and mean numbers of alleles were similar for the Alouette sea-run and resident samples; they were also similar for the Coquitlam sea-run and resident samples.

Analyses of population differentiation using Fst statistics showed no statistically significant pairwise population differentiation between the Alouette sea-run and resident pair, or between the Coquitlam sea-run and resident pair, but all other site pairs were significantly different (Table 2). The genic differentiation test did not show statistically significant differences overall (Fisher’s exact test p = 0.39 for Alouette, and p = 0.09 for Coquitlam), but several individual loci showed statistically significant differentiation (i1, oki1a, oki10, one8, ots2). We did not find a significant isolation by distance pattern (one-tailed p = 0.44) to explain genetic differentiation in terms of distance along waterways between sites.

3.4.2 Genetic bottlenecks
Results from the analyses to identify genetic bottlenecks suggest that there has not been a significant reduction in effective population size for most sites, with the exception of the Alouette reservoir (Table 1). We only detected a significant probability of a bottleneck for the
Chapter 3: Hydroelectric dam induced life-history divergence

Alouette sea-run population (Wilcoxon sign rank $P < 0.05$). Results for rest of the sites suggest that effective population sizes have been relatively stable in the recent past ($P>0.4$).

3.4.3 Population genetic structure: all sites

Analyses to identify genetic clusters in the genotype data provided largely consistent results across different methodologies. The two approaches for determining the optimal number of clusters using the STRUCTURE program (i.e., Likelihood and Evanno methods) provided two different estimates for the number of genetic clusters. The $\text{Ln Pr (X/K)}$ indicated eight genetic clusters ($K=8$) among eleven putative populations (Figure 2a). The $\Delta K$ analysis (Evanno et al. 2005) from the STRUCTURE output suggested that the optimal number of genetic clusters for the data was four ($K=4$; Figure 2b). However, there were high $\Delta K$ values at $K=2$ and $K=8$ as well. The multivariate DAPC analysis suggested that there are eight genetic clusters among the eleven putative populations ($K=8$). Combining the results of the two methodologies, the best estimate appears to be eight genetic clusters.

The assignment of individuals into genetic clusters was remarkably similar for the STRUCTURE and discriminant analysis of principal components (DAPC) methods (Figure 3). At eight genetic clusters, there was no distinction between the Alouette sea-run and resident samples (Figure 3: indicated in red) or the Coquitlam sea-run and resident samples (Figure 3: indicated in purple). Interestingly, Harrison and Pitt River Sockeye sites were also clustered together in one group (mainly blue) and both show admixture with the Big Silver and Birkenhead samples. Although the STRUCTURE analysis separated Harrison and Pitt populations into different clusters at $K=9$, the likelihood of these data given this model is slightly less than the likelihood for $K=8$.

Examination of the STRUCTURE bar plot for $K=4$, which was identified by the Evanno method,
clustered all anadromous Sockeye as one group (cluster 1) except for the Widgeon Slough population (cluster 2), Stave Lake kokanee as one group (cluster 3), and all Alouette and Coquitlam individuals clustered together as the last group (cluster 4). At K=2, which also had a high $\Delta K$ value (Figure 2b), all anadromous sites were separated from the rest of the sites (all of which are considered kokanee, including all of the Alouette and Coquitlam individuals).

3.4.4 Population genetic structure: Alouette and Coquitlam individuals
The DAPC analysis distinguished between Coquitlam sea-run and resident individuals (Figure 4, Figure 5) but no clear distinction was identified between sea-run and resident individuals from the Alouette reservoir (Figure 4). The STRUCTURE results were similar to the DAPC results for the Alouette individuals, but not for the Coquitlam individuals. While the DAPC results showed two clusters for the Coquitlam reservoir, the highest likelihood was for one cluster according to STRUCTURE. However, given that the likelihood support for two genetic clusters in the Coquitlam reservoir was only slightly below the likelihood support of one genetic cluster in the STRUCTURE analysis, and considering the results of both types of analysis, we conclude that there is evidence of weak differentiation between the two types of individuals in the Coquitlam reservoir. This difference was mainly driven by five loci: oki10, oki1a, oki 29, one8, and ots2.

3.4.5 Population splitting time estimate
The mean estimated population splitting time with Isolation with migration (IM) analysis was 2.7 generations (0.4 – 6.8 generations 95% HPD; generation time is 4 years), which is suggestive of very recent splitting of Coquitlam sea-run and resident individuals within the reservoir. This result is inconsistent with the hypothesis that sea-run and resident individuals originated from anadromous and kokanee lineages that existed in the lake prior to dam construction. IM results
Chapter 3: Hydroelectric dam induced life-history divergence

also indicate significantly reduced current population sizes relative to the ancestral population size (Figure 6), consistent with a drastic loss of Sockeye salmon in the Coquitlam basin as a result of the dam.

3.4.6 Long-term historical connectivity vs. contemporary connectivity

The estimates of historical gene flow based on the Migrate-n analysis suggest varying levels of historical genetic connectivity among the eleven Lower Fraser River Sockeye populations. The historical number of immigrants per generation (Nm) ranged from as low as 0.02 (from Harrison to Alouette sea-runs) to 1.99 (from Alouette residents to Alouette sea-runs), a hundred-fold difference (Table 3a). The average number of immigrants per generation was 0.53 for all of the Lower Fraser River populations that we studied. Within the Alouette reservoir, there has been significant gene flow in both directions between sea-run and resident samples, exceeding one migrant per generation (Nm>1). The estimate of gene flow between sea-run and resident populations within the Coquitlam reservoir was lower, yet still moderately high (0.5 <Nm≤1).

Other notable results include the high estimate of historical gene flow into the Birkenhead population from multiple Lower Fraser river populations, mainly from Alouette, Harrison, Pitt River and Weaver Creek (all Nm>1). Other moderately high estimates suggest that there was also considerable historical gene flow into Coquitlam from Alouette (average Nm=0.79), and from both Big Silver and Birkenhead to Weaver Creek (Nm>1).

In contrast to the historical estimates, contemporary gene flow estimates (Nm) were lower for most population pairs (Figure 7, Table 3). As expected, contemporary gene flow from the dammed Alouette and Coquitlam populations to other Lower Fraser River populations was less than historical gene flow in every instance. For example, the average contemporary migration
Chapter 3: Hydroelectric dam induced life-history divergence

(Nm_C) estimate between Alouette and Coquitlam reservoirs is lower than historical estimates (Nm_H) (Nm_C = 0.17 vs. Nm_H = 0.58). Some of the notable declines in contemporary gene flow compared to historical gene flow were observed for immigration into the Birkenhead population from several Lower Fraser River populations (average decline in Nm = 0.9); while historical estimates suggest substantial gene flow into the Birkenhead population from multiple lower Fraser River populations, the contemporary gene flow into the Birkenhead population is substantially less. In contrast, the contemporary gene flow is substantially higher than historical estimates for immigration into the Big Silver population, in this case, from the Birkenhead and Pitt River sites (increase in Nm of 37 and 0.55, respectively). Genetic connectivity among populations measured by gene flow might be different between historical and contemporary time periods due to changes in the migration rate or the effective size of the populations between the two time periods. However, we do not find evidence for changes in effective population size over time for most populations. The bottleneck analysis suggests that the effective population size was reduced significantly in the Alouette reservoir, and the Isolation with Migration (IMa) analysis suggests that there has been a reduction in effective size for Coquitlam reservoir as well. Therefore, differences in historical versus contemporary gene flow among populations, other than from the Alouette or Coquitlam reservoirs, are probably because of changes in migration rate.

3.5 Discussion

3.5.1 Early-stage ecological divergence in sympatry?

Our results indicate that there is a weak genetic distinction between sea-run and resident individuals in the Coquitlam reservoir. Furthermore, our Isolation with Migration (IMa2) analysis suggests that the ecological divergence in life-history types is a very recent
Chapter 3: Hydroelectric dam induced life-history divergence

phenomenon; this bolsters historical observations that the resident life-history type did not exist in Coquitlam Lake before the dam was built. Based on our results and inferences made by Godbout et al. (2011), it is reasonable to infer that some of the anadromous Sockeye salmon that spawned in the lake prior to the construction of the Coquitlam dam were trapped by the dam, and that those have given rise to the two forms of individuals present in the reservoir today. Our analyses suggest that these two forms are in very early stages of divergence within the reservoir, with some individuals still retaining the ability to migrate to the ocean (i.e., sea-run form) and others losing their ability to migrate (i.e., resident kokanee form). The Isolation with Migration analysis suggests that this is a very recent process, meaning that most of the presumed “kokanee” likely had the ability to migrate, if given a chance. In contrast, we did not find a clear distinction between sea-run and resident individuals from the Alouette reservoir.

It is remarkable that any genetic distinction was detected between the two sets of samples from the Coquitlam reservoir, considering the relatively short evolutionary time over which this has occurred after the hydroelectric dam was constructed (~ 100 years since the completion of the dam, ~ 20-25 Sockeye generations) and the absence of a physical barrier to prevent gene flow within the reservoir. Indeed, the presence of two life-history types within each reservoir was not suspected until the chance observations of some salmon migrating downstream towards the ocean during experimental water release and of their subsequent return two years later. We have two lines of evidence that the two forms in the Coquitlam reservoir were derived recently. First, the population splitting time estimates of Isolation with Migration (IM) analysis is consistent with the recent divergence hypothesis, and not consistent with the hypothesis that sea-run and resident forms, currently occupying the reservoir existed as separate historical anadromous and kokanee stocks before the anadromous run was blocked by the dam. This result is also supported
Chapter 3: Hydroelectric dam induced life-history divergence by the absence of historical records of kokanee in Coquitlam Lake before dam construction (Godbout et al. 2011).

Whether or not the two forms in the Coquitlam reservoir will continue to diverge will depend on circumstances in the reservoir. There are three possible fates for these two forms within the reservoir. First, the sea-run form may be lost if the water-release program is not continued and a fish ladder is not constructed to allow access to spawning sites within the reservoir; however, adding these modifications to the reservoir could shape the evolutionary trajectory in two other directions as well. First, if selection for the anadromous form is revived with this access to the ocean and back, it is possible that the sea-run form will displace the resident form entirely. Another possibility is that the sea-run and resident forms may continue to diverge and be maintained as two sympatric forms within the reservoir if there is assortative mating and there is selection against intermediate genotypes. However, although the conversion from migratory to resident life history type has been observed in anadromous fish species in recent history (e.g., Morita et al. 2009), we are not aware of a recent case where both life history types been maintained in sympathy for a long period of time after a recent sympatric divergence event (i.e., occurred during recorded human history).

It is difficult to explain the presence of two forms in the Coquitlam reservoir but not in the Alouette. One possible reason is that Alouette fish have had slightly less time to differentiate than Coquitlam fish (the Coquitlam dam was constructed approximately 14 years earlier than the Alouette). This hypothesis is supported by the population splitting time estimate for Coquitlam, which suggests that splitting took place within the last 20 years. It is also possible that there has been a greater opportunity for the forms to differentiate in the Coquitlam Lake because of a greater diversity of habitats; for example, perhaps there are different types of spawning habitat,
The second goal of this paper was to determine levels of reproductive isolation among Sockeye salmon at 11 sites in the Lower Fraser River. Overall, our results indicate that there are higher levels of genetic connectivity between Sockeye populations than expected, based on previous studies. The literature on Sockeye salmon suggests that this species shows strong natal fidelity and very low levels of “straying” among spawning sites (Hendry et al. 2004; Quinn 1993). It has been thought that this would allow populations to be well adapted to local conditions and differentiate (Taylor 1991; Wood 1995). A recent review by Keefer & Caudill (2014) reported a mean ecological straying rate of 2.4% for Sockeye salmon, but gene flow is expected to be much lower than the ecological straying rate because strays are expected to have much lower reproductive success compared to natives. For example, a recent study by Peterson et al. (2014) found significantly reduced reproductive success for beach-to-stream immigrants where the two sites were less than 1km apart in the Little Togiak Lake, Alaska. In contrast to this finding, which suggests extremely low gene flow among even nearby sites, our results indicate relatively high levels of gene flow between populations that are 100’s of kilometers apart. For example, our estimates of contemporary genetic connectivity between the Pitt River and Harrison site, and between the Pitt River and Big Silver site are unexpectedly high (based on genetic clustering and contemporary gene flow estimates), given that they are not neighboring sites. In addition to gene flow estimates, genetic clustering assignment plots (Figure 3) from both STRUCTURE and
DAPC corroborates the high genetic connectivity among these sites. These apparently high estimates of gene flow cannot be explained by substantial differences in the effective population size between recent and historical times because our analyses do not suggest a substantial change in the effective population size. This leads to the conclusion that the contemporary increase in gene flow is a result of increased migration (probably more straying) between these sites. This could be a result of recent changes in water flow and chemical cues, which would affect rheotaxis and olfaction, both of which are important for orientation in late stages of natal migration (Keefer & Caudill 2014). In particular, further investigation into the reasons for the high level of straying from Pitt River to other populations near Harrison Lake will be of interest for the conservation and management of these populations.

Furthermore, my results in chapter 2 suggest that when population connectivity has declined in recent times, contrasting coalescent-based estimates (e.g., Migrate-n) against disequilibrium-based estimates (e.g., BayesAss) will result in an underestimation of the genetic connectivity change between historical and contemporary times. This is because coalescent-based estimates are likely to underestimate the historical, long-term connectivity, while disequilibrium-based estimates are likely to overestimate contemporary connectivity; this means that the decline in connectivity will be underestimated. Consequently, the change between the two estimators for these Sockeye salmon sites should be thought of as the minimum decline in genetic connectivity. Therefore, it is very likely that genetic connectivity has declined more than our estimates indicate for many lower Fraser River populations. For example, the gene flow decline between Coquitlam and Pitt River is likely to be much higher than the change in our two estimates suggest.
3.5.3 Conservation implications

From a conservation restoration point of view, our findings are encouraging because they suggest that the anadromous form retains its ability to migrate to the ocean, even after a century of being isolated from the ocean. Individuals in both Alouette and Coquitlam reservoirs were presumed to be “kokanee” and yet individuals from both reservoirs did migrate to the ocean and returned two years later. Therefore, in situations where recent barriers (mainly from human infrastructure) are thought to have resulted in a loss of the migratory life-history type, it may be possible to restore the migratory form by facilitating movement through the barrier.

Our results for the eleven Lower Fraser River Sockeye populations we investigated indicate unexpectedly high levels of contemporary gene flow among some of the populations; ecological reasons for this should be investigated and verified. This finding has potential implications for the designation of conservation units and management of these stocks if current genetic connectivity among these sites is high. It will be important to follow up on our result to determine why there has been this recent increase in levels of connectivity; for example, have there been recent changes in the environment at specific sites (e.g., changes in water flow that could affect chemical cues that salmon might use to navigate to their natal spawning site)?

3.5.4 Conclusions

Understanding processes leading to population divergence and eventually to speciation is a long-standing goal in evolutionary biology. In particular, speciation in sympatry has long been a topic of intense debate within the evolutionary biology community (Bird et al. 2012; Mallet et al. 2009; Nosil 2008) and our findings suggest the possibility of very early-stage sympatric ecological divergence within Sockeye salmon. More research is needed to determine the ecological basis for the apparent split into two life-history forms within the Coquitlam reservoir,
but not within the Alouette reservoir. In addition, our molecular analyses of Lower Fraser River Sockeye salmon suggest unexpectedly high levels of recent gene flow between certain sites that are part of different management units under the current delineation of Sockeye management units. Furthermore, the comparison of historical and contemporary migration estimates generally suggests reduced gene flow among most sites, with few notable exceptions. Therefore, sites with substantially elevated levels of recent gene flow should be investigated further to determine ecological reasons for this change and if re-delineating management units are needed to better reflect contemporary population genetic structure and connectivity.

3.6 Acknowledgements

Authors would like to thank Fisheries and Oceans Canada staff at the Pacific Biological Station for generating the data and providing information. In particular, we thank Lyse Godbout for providing background information and providing valuable comments on the manuscript, and John Candy for his assistance in data compilation. We thank Stephen Wright and Ken Minns for providing valuable comments on the manuscript, and Sharcnet High Performance Computer cluster and Grandiflora Computer cluster for computation. Samarasin was supported by NSERC grants (to H. Rodd and B. Shuter) and by the Department of Ecology & Evolutionary Biology at the University of Toronto.
Chapter 3: Hydroelectric dam induced life-history divergence

3.7 Literature cited


Chapter 3: Hydroelectric dam induced life-history divergence


Chapter 3: Hydroelectric dam induced life-history divergence


Chapter 3: Hydroelectric dam induced life-history divergence


Chapter 3: Hydroelectric dam induced life-history divergence


Chapter 3: Hydroelectric dam induced life-history divergence


3.8 Tables

Table 3.1: Measures of genetic diversity for 11 putative Sockeye salmon populations from Lower Fraser River. Abbreviations are as follows: n = sample size, k = mean number of alleles, \( H_o \) = observed gene diversity within individuals, \( H_e \) = expected gene diversity among individuals, \( N_e \) = effective population size estimate (± standard error of the mean), \( P_{Bottleneck} \) = Probability of bottleneck using one-tailed Wilcoxon sign rank test implemented in BOTTLENECK v.1.2.02, assuming 90% stepwise mutations. Bold font indicates a statistically significant probability of a bottleneck.

<table>
<thead>
<tr>
<th>Site/ Population</th>
<th>n</th>
<th>k</th>
<th>( H_o )</th>
<th>( H_e )</th>
<th>( N_e )</th>
<th>( P_{Bottleneck} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alouette sea-runs</td>
<td>60</td>
<td>6.14</td>
<td>0.64</td>
<td>0.64</td>
<td>98.4 (± 2.5)</td>
<td><strong>0.03</strong></td>
</tr>
<tr>
<td>Alouette residents</td>
<td>47</td>
<td>6.71</td>
<td>0.61</td>
<td>0.62</td>
<td>103.8 (± 4.4)</td>
<td>0.45</td>
</tr>
<tr>
<td>Big Silver Creek</td>
<td>46</td>
<td>9.14</td>
<td>0.71</td>
<td>0.71</td>
<td>131 (± 2.7)</td>
<td>0.84</td>
</tr>
<tr>
<td>Birkenhead River</td>
<td>50</td>
<td>9.79</td>
<td>0.71</td>
<td>0.71</td>
<td>147 (± 3.8)</td>
<td>0.97</td>
</tr>
<tr>
<td>Coquitlam sea-runs</td>
<td>48</td>
<td>7.36</td>
<td>0.65</td>
<td>0.68</td>
<td>119 (± 2.9)</td>
<td>0.45</td>
</tr>
<tr>
<td>Coquitlam residents</td>
<td>50</td>
<td>7.50</td>
<td>0.66</td>
<td>0.67</td>
<td>98.6 (± 3.5)</td>
<td>0.68</td>
</tr>
<tr>
<td>Harrison Rapids</td>
<td>42</td>
<td>9.86</td>
<td>0.72</td>
<td>0.71</td>
<td>110.8 (± 0.8)</td>
<td>0.98</td>
</tr>
<tr>
<td>Pitt River</td>
<td>52</td>
<td>10.57</td>
<td>0.71</td>
<td>0.71</td>
<td>126.2 (± 5.2)</td>
<td>0.98</td>
</tr>
<tr>
<td>Weaver Creek</td>
<td>79</td>
<td>10.36</td>
<td>0.68</td>
<td>0.67</td>
<td>102.6 (± 2.0)</td>
<td>0.98</td>
</tr>
<tr>
<td>Widgeon Slough</td>
<td>44</td>
<td>4.29</td>
<td>0.52</td>
<td>0.51</td>
<td>41.8 (± 1.0)</td>
<td>0.86</td>
</tr>
<tr>
<td>Stave Lake</td>
<td>48</td>
<td>8.64</td>
<td>0.61</td>
<td>0.61</td>
<td>108.2 (± 2.1)</td>
<td>0.97</td>
</tr>
</tbody>
</table>
Chapter 3: Hydroelectric dam induced life-history divergence

Table 3.2: Pairwise Fst estimates between 11 putative Sockeye salmon populations estimated using 14 microsatellite loci. Bold values indicate significance (P< 0.00091) after Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th>Alouette sea-runs</th>
<th>Alouette residents</th>
<th>Big Silver Creek</th>
<th>Birkenhead River</th>
<th>Coquitlam sea-runs</th>
<th>Coquitlam residents</th>
<th>Harrison Rapids</th>
<th>Pitt River</th>
<th>Weaver Creek</th>
<th>Widgeon Slough</th>
<th>Stave Lake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alouette sea-runs</td>
<td>-0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alouette residents</td>
<td></td>
<td>-0.002</td>
<td>0.065</td>
<td>0.072</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big Silver Creek</td>
<td>0.065</td>
<td>0.072</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birkenhead River</td>
<td>0.109</td>
<td>0.123</td>
<td>0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coquitlam sea-runs</td>
<td>0.044</td>
<td>0.057</td>
<td>0.057</td>
<td>0.084</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coquitlam residents</td>
<td>0.026</td>
<td>0.036</td>
<td>0.044</td>
<td>0.071</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harrison Rapids</td>
<td>0.076</td>
<td>0.090</td>
<td>0.021</td>
<td>0.027</td>
<td>0.061</td>
<td>0.058</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitt River</td>
<td>0.083</td>
<td>0.093</td>
<td>0.022</td>
<td>0.027</td>
<td>0.060</td>
<td>0.047</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weaver Creek</td>
<td>0.092</td>
<td>0.106</td>
<td>0.026</td>
<td>0.034</td>
<td>0.073</td>
<td>0.060</td>
<td>0.027</td>
<td>0.037</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widgeon Slough</td>
<td>0.193</td>
<td>0.213</td>
<td>0.153</td>
<td>0.130</td>
<td>0.208</td>
<td>0.189</td>
<td>0.177</td>
<td>0.143</td>
<td>0.161</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stave Lake</td>
<td>0.147</td>
<td>0.159</td>
<td>0.104</td>
<td>0.108</td>
<td>0.105</td>
<td>0.086</td>
<td>0.121</td>
<td>0.082</td>
<td>0.143</td>
<td>0.294</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3: Hydroelectric dam induced life-history divergence

Table 3.3: Estimates of gene flow ($N_{m}$) among lower Fraser River Sockeye salmon populations, based on a) Migrate-n analysis, and b) BayesAss analysis.

<table>
<thead>
<tr>
<th>FROM</th>
<th>TO</th>
<th>Alouette sea-runs</th>
<th>Alouette residents</th>
<th>Big Silver Creek</th>
<th>Birkenhead River</th>
<th>Coquitlam sea-runs</th>
<th>Coquitlam residents</th>
<th>Harrison Rapids</th>
<th>Pitt River</th>
<th>Weaver Creek</th>
<th>Widgeon Slough</th>
<th>Stave Lake</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alouette sea-runs</td>
<td>---</td>
<td>1.15</td>
<td>0.50</td>
<td>0.86</td>
<td>0.69</td>
<td>0.81</td>
<td>0.60</td>
<td>0.28</td>
<td>0.70</td>
<td>0.18</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>Alouette residents</td>
<td>1.99</td>
<td>---</td>
<td>0.92</td>
<td>1.24</td>
<td>0.55</td>
<td>1.06</td>
<td>0.73</td>
<td>0.39</td>
<td>0.78</td>
<td>0.46</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Big Silver Creek</td>
<td>0.27</td>
<td>0.24</td>
<td>---</td>
<td>1.09</td>
<td>0.24</td>
<td>0.52</td>
<td>0.76</td>
<td>0.48</td>
<td>1.20</td>
<td>0.11</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>Birkenhead River</td>
<td>0.22</td>
<td>0.23</td>
<td>0.81</td>
<td>---</td>
<td>0.26</td>
<td>0.40</td>
<td>0.69</td>
<td>0.54</td>
<td>1.02</td>
<td>0.16</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Coquitlam sea-runs</td>
<td>0.43</td>
<td>0.33</td>
<td>0.50</td>
<td>0.90</td>
<td>---</td>
<td>1.00</td>
<td>0.55</td>
<td>0.31</td>
<td>0.54</td>
<td>0.12</td>
<td>0.29</td>
<td></td>
</tr>
<tr>
<td>Coquitlam residents</td>
<td>0.49</td>
<td>0.26</td>
<td>0.57</td>
<td>0.58</td>
<td>0.61</td>
<td>---</td>
<td>0.53</td>
<td>0.37</td>
<td>0.62</td>
<td>0.11</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Harrison Rapids</td>
<td>0.02</td>
<td>0.19</td>
<td>0.59</td>
<td>1.09</td>
<td>0.16</td>
<td>0.37</td>
<td>---</td>
<td>0.39</td>
<td>0.54</td>
<td>0.11</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>Pitt River</td>
<td>0.31</td>
<td>0.27</td>
<td>0.66</td>
<td>1.15</td>
<td>0.19</td>
<td>0.52</td>
<td>0.63</td>
<td>---</td>
<td>0.91</td>
<td>0.17</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Weaver Creek</td>
<td>0.43</td>
<td>0.30</td>
<td>1.30</td>
<td>1.97</td>
<td>0.30</td>
<td>0.51</td>
<td>0.80</td>
<td>0.43</td>
<td>---</td>
<td>0.20</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Widgeon Slough</td>
<td>0.24</td>
<td>0.26</td>
<td>0.43</td>
<td>0.77</td>
<td>0.16</td>
<td>0.29</td>
<td>0.61</td>
<td>0.39</td>
<td>0.47</td>
<td>---</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Stave Lake</td>
<td>0.18</td>
<td>0.25</td>
<td>0.83</td>
<td>0.65</td>
<td>0.27</td>
<td>0.41</td>
<td>0.60</td>
<td>0.38</td>
<td>0.89</td>
<td>0.15</td>
<td>---</td>
<td></td>
</tr>
</tbody>
</table>
### Chapter 3: Hydroelectric dam induced life-history divergence

#### b)

<table>
<thead>
<tr>
<th></th>
<th>Alouette sea-runs</th>
<th>Alouette residents</th>
<th>Big Silver Creek</th>
<th>Birkenhead River</th>
<th>Coquitlam sea-runs</th>
<th>Coquitlam residents</th>
<th>Harrison Rapids</th>
<th>Pitt River</th>
<th>Weaver Creek</th>
<th>Widgeon Slough</th>
<th>Stave Lake</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alouette sea-run</strong></td>
<td>---</td>
<td>31.58</td>
<td>0.29</td>
<td>0.11</td>
<td>0.39</td>
<td>0.08</td>
<td>0.31</td>
<td>0.10</td>
<td>0.12</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Alouette resident</strong></td>
<td>0.05</td>
<td>---</td>
<td>0.29</td>
<td>0.11</td>
<td>0.27</td>
<td>0.06</td>
<td>0.27</td>
<td>0.09</td>
<td>0.07</td>
<td>0.03</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Big Silver Creek</strong></td>
<td>0.06</td>
<td>0.24</td>
<td>---</td>
<td>0.12</td>
<td>0.27</td>
<td>0.07</td>
<td>0.29</td>
<td>0.09</td>
<td>0.07</td>
<td>0.03</td>
<td>0.07</td>
</tr>
<tr>
<td><strong>Birkenhead River</strong></td>
<td>0.06</td>
<td>0.24</td>
<td>38.21</td>
<td>---</td>
<td>0.26</td>
<td>0.07</td>
<td>1.00</td>
<td>0.12</td>
<td>0.84</td>
<td>0.03</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>FROM</strong></td>
<td>Coquitlam sea-runs</td>
<td>0.05</td>
<td>0.24</td>
<td>0.29</td>
<td>0.11</td>
<td>---</td>
<td>0.06</td>
<td>0.28</td>
<td>0.09</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Coquitlam residents</td>
<td>0.06</td>
<td>0.25</td>
<td>0.30</td>
<td>0.13</td>
<td>36.32</td>
<td>---</td>
<td>0.34</td>
<td>0.10</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Harrison</td>
<td>0.05</td>
<td>0.24</td>
<td>0.30</td>
<td>0.11</td>
<td>0.27</td>
<td>0.07</td>
<td>---</td>
<td>0.09</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Pitt River</td>
<td>0.05</td>
<td>0.24</td>
<td>1.21</td>
<td>0.17</td>
<td>0.27</td>
<td>0.07</td>
<td>32.49</td>
<td>---</td>
<td>0.15</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Weaver Creek</td>
<td>0.05</td>
<td>0.24</td>
<td>1.28</td>
<td>0.15</td>
<td>0.28</td>
<td>0.07</td>
<td>0.55</td>
<td>0.11</td>
<td>---</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Widgeon Slough</td>
<td>0.06</td>
<td>0.25</td>
<td>0.29</td>
<td>0.12</td>
<td>0.28</td>
<td>0.07</td>
<td>0.28</td>
<td>0.10</td>
<td>---</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Stave Lake</td>
<td>0.05</td>
<td>0.24</td>
<td>0.30</td>
<td>0.12</td>
<td>0.27</td>
<td>0.07</td>
<td>0.28</td>
<td>0.12</td>
<td>0.07</td>
<td>0.03</td>
</tr>
</tbody>
</table>
3.9 Figures

Figure 3.1: Schematic of the study area and Sockeye Salmon sites that were studied.
Figure 3.2: a) Probability of data as a function of K (the number of clusters) in STRUCTURE analysis suggest K=8 is the optimal number of clusters. b) Changes of $\Delta K$ as a function of K (the number of clusters) suggest K=4 as the optimal number of clusters according to Evanno method
Figure 3.3: Assignment of individuals to genetic clusters using a) Bayesian clustering program STRUCTURE and b) Discriminant analysis of principal components (DAPC) method at optimal genetic clusters (K=8). Both methods show very similar pattern of individual assignments to genetic clusters.
Figure 3.4: Discriminant analysis of principal components (DAPC) scatter plot with only Alouette and Coquitlam individuals.
Chapter 3: Hydroelectric dam induced life-history divergence

Figure 3.5: Bar plot showing the assignment of individuals (sea-run and resident) from the Coquitlam reservoir into two genetic clusters (K=2) using DAPC method show a clear distinction in the genetic makeup between sea-run and resident individuals.
Figure 3.6: Results of IM analysis showing estimated population sizes, and population splitting time. Time is represented in the vertical axis. Widths of boxes are representative of current and ancestral population size estimates, and horizontal line represent population splitting time (95 % highest posterior density (HPD) in dashed horizontal lines).
Chapter 3: Hydroelectric dam induced life-history divergence

Figure 3.7: Difference between historical and contemporary gene flow ($N_{mH} - N_{mC}$) estimates based on Migrate-n and BayesAss analysis. Positive values indicate higher historical gene flow relative to contemporary gene flow.
Chapter 4

Fish diversity and biomass in northern Canadian lakes: northern lakes are more diverse and have greater biomass than expected based on species-energy theory

4.1 Abstract

Biodiversity in northern Canada (north of 60º N latitude) is threatened, primarily by increasing resource exploitation and by climate change. Unfortunately, we have relatively limited knowledge of aquatic biodiversity for this region, making it difficult to develop suitable policies to manage these threats. Here, we describe, quantify, and test hypotheses related to fish biodiversity and biomass in 37 lakes in a diamond-mining district (the Barrenlands) in the Northwest Territories, Canada (64ºN 110ºW). To estimate species richness and biomass of fishes, we took advantage of exhaustive sampling and monitoring surveys conducted in the region and compared our northern estimates against estimates from southern Canadian lakes. We found that most of the 37 northern lakes contained 2-4 species with the largest lake containing 8 species. Salmonids dominated this system, with lake trout (Salvelinus namaycush) being the dominant species in abundance and biomass. Comparative analysis with similar-sized southern Canadian lakes showed no significant difference in the slopes of species richness versus lake area curves. Surprisingly, total fish biomass distributions for northern, Barrenlands lakes were also similar to southern, Ontario lakes. Overall, our results suggest that Barrenlands lakes are important natural resources of Canada that should be conserved for the future. Under anticipated scenarios of climate change, these lakes may represent important refugia for coldwater fishes (e.g. lake trout) as habitats at the southern edges of their ranges become more limiting.
Chapter 4: Fish diversity and biomass in northern Canadian lakes

4.2 Introduction

Species richness (number of species in a given area) is a fundamental measure of diversity (MacArthur and Wilson 1967), and explaining patterns of diversity is a long-standing goal of ecology. At the global scale, biodiversity and productivity generally decline when moving poleward from the equator (Rosenzweig 1995; Kaufman and Willig 1998; Hillebrand 2004; Lewis 2011), a pattern that includes freshwater fishes (Barbour and Brown 1974; Oberdorff et al. 1995). Declines in species diversity at higher latitudes have also been observed at smaller regional scales. For example, Mandrak (1995) found a south-to-north decline in fish species richness for Ontario (Canada) lakes. This finding, if extrapolated, suggests that northern lakes (defined here as lakes north of 60° N latitude) should have significantly lower species richness relative to lakes in temperate regions. However, difficulties in conducting field work and the general expectation of low fish diversity and productivity in northern regions have resulted in a limited scientific literature on fish in this region compared to more temperate latitudes. For example, Downing and Plante (1993) analyzed fish production from 38 lakes globally and their dataset included only one lake from northern Canada (Char Lake, Nunavut). These gaps in knowledge may create the perception that the value of northern fishes is low. In turn, increasing resource development in northern Canada could be rationalized by this perception of low biodiversity value, and the harmful alteration or destruction of fish habitat that often accompanies mining activity could accelerate biodiversity loss in the north (Chu et al., 2015).

Here, we investigate fish diversity and biomass of northern lakes in a comparative framework to test hypotheses related to diversity and biomass. Based on species-energy theory (Wright 1983; Currie 1991; Evans et al. 2005) and previously described latitudinal patterns (Mandrak 1995), we expected fish diversity, abundance, and biomass to decrease with decreasing energy availability.
Chapter 4: Fish diversity and biomass in northern Canadian lakes

(i.e., increasing latitude). Energy in a system is typically measured as available solar radiation (Johnson 1994) or using a normalized vegetation index (Storch et al. 2005), depending on the system of interest. Both long term regional processes such as historical biogeography and climate (Jackson and Harvey 1989; Mandrak 1995; Griffiths 2010), as well as local factors such as lake area and water chemistry (Matuszek and Beggs 1988; Allen et al. 1999), have been shown to explain variation in species richness in temperate regions and it would be useful to evaluate the relative importance of these factors in northern regions.

Although species richness increases with area (Arrhenius 1921; Gleason 1922) and with energy (Wright 1983; Currie 1991), there appears to be a negative interaction between area and energy (Storch et al. 2005). According to Storch et al. (2005), this interaction causes the slope of the log-transformed species-area relationship to be lower in areas with high energy compared to areas with low energy. That is, the exponent \( z \) in the power equation \( \text{Species Richness} = c \cdot \text{Area}^z \) is expected to be smaller for high energy areas compared to low energy areas. This pattern has been found in empirical studies by Lyons and Willig (2002) and Storch et al. (2005). In contrast, the intercept of the log-transformed species- area relationship (\( \log[c] \)) is expected to be elevated for high energy areas, and the overall relationship for high energy areas is expected to be elevated above the low energy areas. Several explanations have been put forth to explain this phenomenon including: 1) higher rates of speciation in high energy areas, 2) lower rates of extinction in high energy areas, and 3) the physiological tolerance hypothesis, which predicts that more species can tolerate warm and wet conditions compared to cold and dry conditions. However there has not been strong support for any particular hypothesis (reviewed by Currie et al. 2004). Nevertheless, based on the observed empirical patterns, we expect the slope of the
Chapter 4: Fish diversity and biomass in northern Canadian lakes

species-area relationship for southern Canadian lakes to be lower than the slope for northern Canadian lakes, but the intercept is expected to be higher for southern lakes.

Similar to species richness, we expect total fish biomass to decrease with increasing latitude because primary production decreases with increasing latitude (Liu et al. 2002; Lewis 2011) and this reduction in production should exert a bottom-up constraint on consumers such as fishes. Downing and Plante (1993) found a positive relationship between annual fish production (kg ha\(^{-1}\) yr\(^{-1}\)) and mean annual standing fish biomass (kg ha\(^{-1}\)). Their analysis also showed positive correlations between fish production, temperature, and primary production. Because of these relationships, we predict that the fish biomass of northern lakes will be significantly lower than that of temperate-zone lakes.

We compared fish diversity and resources (biomass was used as a proxy) from a northern Canadian region (the Barrenlands; Figure 1) against similar-sized lakes in southern Canada (primarily Ontario). We predicted that log-transformed species richness would increase at a higher rate with lake area (i.e., have steeper slope) in Barrenlands lakes compared to southern Canadian lakes because of the expected negative interaction between area and energy (Storch et al. 2005). We also predicted that variation in species richness among regions would be largely due to regional factors (historical biogeography and climate) rather than local factors (e.g., lake area) based on previous research on similar aquatic systems (Tonn 1990; Beisner et al. 2006). We also explored the relationships of fish biomass with lake area, mean annual air temperature, mean depth, and number of species to identify which variables were associated with differences in fish biomass over a broad regional scale. Air temperature was used as a proxy for lake water temperature. Air temperature has been found to be a good predictor of lake water temperature by Sharma et al. (2008), and has been used in models to predict water temperature by others.
Chapter 4: Fish diversity and biomass in northern Canadian lakes
(Matuszek and Shuter 1996; Trumpickas et al. 2009). Finally, we estimated the abundance and biomass of dominant species and described patterns we ascertained with respect to biomass. We expected coldwater species to dominate Barrenlands lakes in numerical abundance and biomass, but, because of energetic constraints, we predicted that the biomass of the dominant coldwater species would be lower than in temperate lakes (Wright 1983).

4.3 Materials and methods

4.3.1 Study area and data

The study lakes are located in a diamond-mining district in the Barrenlands region (Figure 1) of the Northwest Territories, Canada, about 330 km northeast of Yellowknife (approximately 64° 45’ N, 110° 38’ W). The mining companies operating in this region were required to collect fish and habitat data from local lakes by Fisheries and Oceans Canada (DFO). In some lakes, all the fishes were removed and the lake was drained for mine development (to be referred to as “fish-out lakes” from here on; Tyson et al. 2011), while other lakes were sampled for monitoring purposes (“monitored lakes” from here on). The data collected included habitat variables such as lake surface area, depth, and water quality; fish data such as species caught, fork length, and weight of individual fish; and fishing effort data such as fished times and gear used. Most lakes had an area of 20 -200 ha. Fish were sampled predominantly using gillnets with mesh sizes ranging from 0.7 cm to 14 cm (0.7, 2.2, 3.8, 5.1, 6.4, 7.6, 10.2, 12.7, and 14 cm), including multi-mesh nets. Trap nets and angling were used in a few lakes but this effort was minimal relative to gillnet effort. The database contained 53 lakes or ponds. We used all 53 lakes and ponds for the preliminary regional species richness analysis to document species occurrence in the Barrenlands region (Analysis a in Table 1). Due to differences in sampling intensity among lakes, we applied certain inclusion criteria to select lakes that were adequately sampled for
subsequent species richness analyses. These criteria are specified in subsequent sections in the Methods and summarized in Table 1. Comparative data for the regional analysis of the relationship between lake area and species richness were obtained from study sites across Canada, as reported in the following sources: Black-Hollow river region, ON (Jackson 1988; Jackson and Harvey 1989); Bruce Peninsula, ON (Harvey 1981); Manitoulin Islands, ON (Harvey 1978); Wawa region, ON (Somers 1980; Somers and Harvey 1984); Athabasca region, AB (Robinson and Tonn 1989); Vilas County, WI (Tonn and Magnuson 1982). Fish species richness for secondary watersheds were obtained by pooling tertiary watershed species lists in Chu et al. (2003) according to described secondary watersheds in Minns et al. (2008).

For the comparative analysis of biomass, fish biomass data from lakes were collated from the DFO Barrenlands data set and the published literature, along with variables such as lake latitude/longitude, lake area, mean depth, and number of species (Appendix 1). To estimate the mean annual air temperature for lakes, we used 0.5 degree gridded 1961–1990 temperature data from the CRU TS 2.1 dataset (Mitchell and Jones 2005) available at http://www.ipcc-data.org/obs/cru_ts2_1.html. We obtained lake trout (Salvelinus namaycush) biomass data for Canadian lakes, through the Ontario Ministry of Natural Resources (Lester 2013). The data are mainly from Ontario lakes, but include lakes from Quebec, Yukon, Alberta, and Saskatchewan. Table 1 summarizes the analyses conducted and the datasets used.

4.3.2 Analyses

All calculations, statistical analyses, and plots were prepared using R version 2.14.2 (R Development Core Team 2012) unless stated otherwise.
4.3.2.1 Species richness

Species richness for each monitored lake was estimated using EstimateS 8.2.0 (Colwell 2006), which uses sample- or individual-based rarefaction or extrapolation curves to estimate the “actual” species richness. First, we plotted all species richness estimators implemented in the program EstimateS (e.g., Chao, ICE, Jackknife, Michaelis-Menten) as a function of number of sampled individuals in fish-out lakes where we assumed we knew the “actual” species richness because the sampling was exhaustive. We found the species richness estimates from different methods to be very similar for most fish-out lakes, but the Jackknife 1 estimator (Heltshe and Forrester 1983) provided the most accurate estimate of the actual species richness with the least amount of sampling. The Jackknife 1 method estimates total species richness, including species not present in any sample, by incorporating a term for unsampled species in the system, the magnitude of which decreases with increasing sampling effort (Burnham and Overton 1978; Heltshe and Forrester 1983). It provided the best estimate of actual species richness of fish-out lakes when at least 95 individuals were sampled. Therefore, Jackknife 1 was selected to estimate species richness in monitoring lakes that had at least 95 sampled individuals (Analysis b in Table 1). The estimated species richness did not differ from observed species richness in most of the lakes that had at least 95 sampled individuals; in fact, the estimated number of species using the Jackknife 1 method, rounded to the nearest whole number, was identical to the observed species richness in 15 of 18 monitored lakes. For the other three monitored lakes, one had an estimate higher by one species (+1) and two lakes had estimates higher by two species (+2).

We compared the relationships between species richness and lake area among different regions by testing if slopes and intercepts of species–area regression lines were significantly different. Analysis of covariance (ANCOVA) was used to test for heterogeneity in slopes by including the
Chapter 4: Fish diversity and biomass in northern Canadian lakes

region as the categorical factor (Analysis c in Table 1). Lake area (in hectares) and species richness were log_{10}-transformed to stabilize variance so that the assumption of normality for the residuals was met. Multiple regression analysis was conducted to explain species richness in terms of local (lake area) and regional (secondary watershed species richness) factors (Analysis d in Table 1). The best explanatory model was chosen based on corrected Akaike information criteria (AICc: Hurvich and Tsai 1989; Anderson and Burnham 2002).

4.3.2.2 Fish abundance

Fishing effort and catches were calculated from sampling data. For fish-out lakes, we used maximum likelihood and regression-based catch-effort models to estimate total fish abundance. We first estimated abundance using standard estimators including the Leslie-Ricker (Leslie and Davis 1939; Ricker 1975), DeLury-Ricker (Delury 1947; Ricker 1975) and maximum-likelihood (Gould and Pollock 1997) methods. Generally, estimates from different methods were consistent for most lakes. We found that the Leslie-Ricker estimates for fish-out lakes were closest to the total catches obtained from exhaustive sampling in fish-out lakes. Because few of our lakes were sampled with multiple gear types, we needed to modify the Leslie-Ricker method to account for differences in fishing effort resulting from different gear types. We regressed gillnet catch-per-effort (dependent variable) against the cumulative catch from all gears (independent variable) and calculated the initial abundance by dividing the intercept by the negative slope of the regression (i.e., total cumulative catch when catch-per-effort is zero). This is identical to the Leslie-Ricker estimation of abundance, only differing by the use of cumulative catch from all gear types, instead of just one. This method yielded abundance estimates identical to the Leslie-Ricker estimates when all the fishing effort was from one gear type, and provided more accurate estimates of total abundance in lakes where multiple gear types were used. Catch-per-effort
models assume that enough fish were removed to substantially reduce catch per unit effort over
time (Leslie and Davis 1939), which may not be true for monitored lakes where sampling effort
was much lower than fish-out lakes. Therefore, we used regression analyses of the fish-out data
set to assess the feasibility of using initial catch per effort (CPE₀) values to estimate total
abundance in monitored lakes (Analysis e in Table 1).

4.3.2.3  Biomass and dominance

Fish biomass for each lake was estimated by using the estimated fish abundance and average
weight of fishes caught in the lake. To estimate the biomass of individual species, we determined
the fraction of fishes belonging to each species from catch data. This was multiplied by the total
fish abundance to calculate the abundance of the particular species and then multiplied by the
average weight of the species in the lake.

We used the Wilcoxon rank sum test to compare fish biomass distributions from Ontario (Kelso
and Johnson 1991) and Barrenlands lakes (Analysis f in Table 1). Using a larger dataset
(Appendix 1) that included lakes from a broad geographic area, we analyzed relationships
between fish biomass, lake area, mean depth, mean annual air temperature, and number of
species using multiple regression after log₁₀-transformation of fish biomass, lake area, and mean
depth (Analysis g in Table 1). The model best able to account for observed variation in fish
biomass was chosen based on corrected Akaike information criteria (AICc). The contributions of
the dominant species to total abundance and biomass were quantified and the relative abundance
and biomass of the dominant species were plotted as functions of number of species in the lake.
The relationship between lake trout biomass, lake area, and latitude was evaluated using multiple
regression after log₁₀-transformations of biomass and lake area (Analysis h in Table 1).
Chapter 4: Fish diversity and biomass in northern Canadian lakes

4.4 Results

Nine fish species were found in our Barrenlands lakes. lake trout, Arctic grayling (*Thymallus arcticus*) and round whitefish (*Prosopium cylindraceum*) were the most common species, inhabiting at least 32 of the 53 lakes (Figure 2). Most lakes contained 2-4 species, but the largest lake (Long Lake) contained 8 species. We found significantly more species in fish-out lakes relative to monitored lakes, presumably because of the incomplete sampling of the latter (Wilcoxon rank sum test $W=137$, 1-tailed $p=0.01$).

4.4.1 Species – area relationship

Log$_{10}$ (species richness) was linearly related to the log$_{10}$ (lake area) in fish-out lakes (Figure 3: $R^2=0.53$, $F_{1,8}=7.95$, $p=0.026$), but the strength of the association (judged by the $R^2$ value) weakened when monitored lakes were included ($R^2=0.22$, $F_{1,25}=6.94$, $p=0.014$). Relationships between species richness and the lake area among different regions showed no significant difference in slopes (ANCOVA: lake area by region interaction $F_{1,6}=0.98$, $p=0.436$; Figure 4). However, there were significant differences in the intercepts ($F_{1,6}=43.12$, $p<<0.001$), which ranged from 0.05 for the Barrenlands lakes to 0.64 for Wawa, Ontario. Overall, for a 100 ha lake, average species richness was 3.7 species in Barrenland lakes, whereas species richness ranged from 2.6 in Alberta (lowest) to 11.2 in Bruce Peninsula, Ontario (highest).

According to AICc, the additive multiple regression model (equation 1) that included both lake area and secondary watershed species richness was best able to account for the observed variation in lake species richness, with secondary watershed species richness accounting for 34% of the observed variance and lake area accounting for 22%.
Chapter 4: Fish diversity and biomass in northern Canadian lakes

(1) \[ \log_{10} \text{Lake species richness} = -0.73 + 0.24 \log_{10} \text{Lake area} + 0.62 \log_{10} \text{Secondary watershed species} \]

\[ (R^2 = 0.56, F_{2, 279} = 177.8, p << 0.001). \]

4.4.2 Fish abundance and total biomass

\[ \log_{10}\text{-transformed initial catch per effort (CPE)}_0 \text{ was linearly related to fish density (in number per hectare; } R^2 = 0.59, p = 0.005) \text{ and total fish abundance } (R^2 = 0.79, p < 0.001) \text{ for the fish-out lakes. Therefore, we used initial catch per effort of monitored lakes to better estimate the total abundance and biomass of monitored lakes. Total fish biomass in Barrenlands lakes ranged from 0.2 to 25 kg·ha}^{-1}, \text{ with a median of } 10.1 \text{ kg·ha}^{-1}. \text{ Total fish biomass distribution in Barrenland lakes marginally differed from the small Ontario lakes (Wilcoxon rank sum test } W = 170, p = 0.08; \text{ Figure 5). Our multiple regression analysis using lakes from a broad geographic scale (Appendix 1) showed that observed variation in total fish biomass was best accounted for by differences in lake area, mean annual air temperature, and number of species in the lake:} \]

(2) \[ \log_{10} \text{Biomass} = 2.49 + 1.04 (\log_{10} \text{Lake area}) + 0.11 (\text{Number of species}) + 0.08 (\text{Mean annual air temperature}) \]

\[ (R^2 = 0.77, F_{3, 52} = 59.06, p < 0.001). \]
Chapter 4: Fish diversity and biomass in northern Canadian lakes

Log-transformed total fish biomass per lake was strongly associated with log$_{10}$ lake area (Figure 6a) and the coefficient (1.04) was not significantly different from one. Biomass (kg. ha$^{-1}$) was also positively associated with species richness and mean annual air temperature (Figure 6b).

4.4.3 Dominant species

In terms of abundance, Barrenlands lakes were largely dominated by either round whitefish (14 of 35 lakes) or lake trout (12 of 35 lakes). Lake trout, however, had the greatest biomass in 18 out of 31 lakes. The dominant species contributed a substantial proportion to the total fish abundance and biomass in lakes with more than one species (Figure 7). Percent abundance of the dominant species declined as species richness increased (Figure 7a), but still made up 38% of the fish present in the lake with 8 species. Similarly, the biomass contribution of the dominant species decreased with increasing species richness, but rarely fell below 50% (Figure 7b). The dominant species’ contribution to total fish biomass in the Barrenlands lakes was similar to that found in Ontario lakes (Figure 7b).

Multiple regression analysis showed no significant effect of latitude on lake trout biomass ($p = 0.75$) but lake area was significant ($p << 0.001$ $R^2 = 0.91$, $F_{2,54} = 262.2$). The slope of the regression line (1.13) was significantly different from 1 ($p = 0.029$). Even if biomass was expressed as kg/ha (i.e. assuming a 1:1 relationship between area and biomass), there was no significant linear relationship between lake trout biomass and latitude (Figure 8: $R^2 = 0.01$, $F_{1,55} = 0.6$, $p = 0.44$).
4.5 Discussion

Our results indicate that freshwater fish diversity and resources in northern lakes are greater than the general perception and are, in fact, comparable to regions in southern Canada. We found that small lakes in the northern, Barrenlands region contained up to 8 species per lake and were dominated by coldwater species such as lake trout, round whitefish, and Arctic grayling.

Contrary to expectations from the literature, the slopes of the relationships linking species richness to lake area were not significantly different between northern, Barrenlands lakes and southern, Ontario lakes. Furthermore, our comparative analysis suggests that regional factors are an important determinant of lake fish assemblages.

Contrary to our expectations, we found the Barrenlands fish biomass distribution to be similar to that found in small Ontario lakes. Mandrak (1995) found a gradient in species richness in Ontario, with richness declining from south to north; this led us to expect that Barrenlands lakes, in the far north, would have far fewer species than the lakes in Ontario. However, Mandrak’s regional analysis was done at the quadrat level and did not explicitly consider lake area, which often explains considerable variation in species richness at a local scale (e.g., Matuszek and Beggs 1988; Allen et al. 1999). Our comparative analysis, which accounted for lake area, showed that species richness increased with lake area in a similar fashion for both Barrenlands and southern lakes. This result of no significant difference in slopes of species-area curves does not support Storch et al.’s (2005) prediction that low energy areas would have a higher slope than high energy areas and thus questions the generality of a negative interaction between area and energy with respect to species richness. It is possible that the negative interaction is more apparent when two areas with large differences in energy are contrasted (as in the case for South
Africa versus Britain; Storch et al. 2005), but is less apparent when the differences in energy are moderate (Barrenlands vs. Ontario).

Although the slope of the species–area relationship was similar across regions, we did find that the northern lakes contained fewer species per unit lake area than southern lakes, as shown by a significantly lower intercept for the Barrenlands species-area curve. The intercept of the species-area curve can be thought of as representing the regional species pool, where high regional richness would increase the intercept because individual lakes are more likely to be colonized by more species. This is consistent with Lyons and Willig’s (2002) finding of higher intercept values for the tropics compared to temperate regions. Regional species pools in lower latitudes tend to be higher because of higher energetics and longer biogeographic history; northern Canadian lakes are younger and farther from glacial refugia, both of which reduce the number of species available to colonize lakes because of dispersal limitations and less time for speciation to occur (Bernatchez and Wilson 1998; Shafer et al. 2010). For example, Tonn et al. (1990) compared regional fish assemblages in two areas with similar climatic variables (Wisconsin, USA versus Finland) and attributed most of the higher species richness in Wisconsin lakes to differences in biogeographic histories. Therefore, we infer that the reduced species richness in Barrenlands lakes is largely driven by smaller species pools for this region. This inference is further supported by our analysis that showed a regional factor (i.e., secondary watershed species richness) was a better predictor of lake species richness than a local factor (i.e., lake area). The relative importance of secondary watershed species richness suggests that long term regional processes, such as historical biogeography and dispersal, play significant roles in determining local diversity (Tonn et al. 1990; Beisner et al. 2006; see also Angermeier and Winston 1998; Niu et al. 2012).
4.5.1 Biomass

Biomass is a key ecological variable for quantifying fish resources in a system. Our estimates of fish biomass for the northern lakes ranged from 0.2 to 25 kg.ha\(^{-1}\). Although we expected southern lakes to have higher biomass levels for energetic reasons (Wright 1983; Evans et al. 2005), we found only a marginal difference between the biomass distributions of small Ontario lakes and Barrenlands lakes. Fish biomass estimates for high latitude lakes are rare in the scientific literature (MacCallum and Regier 1984), but the Malinen et al. (2013) estimates of European whitefish biomass (\textit{Coregonus lavaretus}) in six subarctic Finnish lakes provide a useful comparison for our study because whitefishes are dominant in both regions. Their estimates of biomass cover a range (0.5 – 13.3 kg.ha\(^{-1}\)) that is similar to our estimates for the Barrenlands lakes, even though mean annual air temperature of Barrenland lakes is 5-8 °C lower than these Finnish lakes. This suggests that northern Canadian lakes contain more fish biomass than might be expected based on temperature.

Our analysis of lake fish biomass over a broad geographic range indicated that total biomass increased with lake area, mean annual air temperature, and species richness of the lake. The positive association between species richness and biomass could be due to effects of niche complementarity (Tilman et al. 1997), where more complex systems allow the coexistence of more species (through niche partitioning), and this results in higher total biomass. Similar to Cote et al. (2011), we did not find lake depth to be associated with differences in biomass. This contrasts with previous studies that identified lake depth as an important predictor (Matuszek 1978; Hanson and Leggett 1982). Although the lakes analyzed by Cote et al. (2011) are relatively small, shallow lakes, our analysis contained lakes that covered a broad range in depths, in addition to the Barrenlands lakes (see Appendix 1).
Biomass and production are key ecological variables that help to ascertain the value of fish resources in a system. Although the relationship between fish production and biomass over a broad geographic scale was found to be close to one-to-one by Downing and Plante (1993), we do not expect the 1:1 relationship to hold in northern lakes with cold waters. Northern Canadian lakes, including Barrenland lakes, are occupied by long-living, slow growing fish with relatively large size (Johnson 1976; Sparholt 1985). Because production to biomass ratios are inversely related to maximum size and positively related to lake productivity, we expect lower production to biomass ratios in northern Canadian lakes (Randall and Minns 2000; Giacomini et al. 2013). Therefore, although biomass might be relatively similar in northern and southern lakes, we expect overall production to be lower. The observation of low recruitment in arctic lakes, which are dominated by salmonids such as lake trout, and Arctic charr (Johnson 1994) lends further support to this expectation.

4.5.2 Dominant species

Coldwater species in the Salmonidae family, especially round whitefish and/or lake trout, dominated Barrenlands lakes. In the northern Nordic region, which has a similar latitude to the Barrenlands, Percidae (Perch), Esocidae (Pike), together with Salmonidae, are most common (Lehtonen et al. 2008), likely because Nordic lakes are warmer than subarctic North American lakes (Shuter et al. 2012). In contrast, small Ontario lakes are dominated by a combination of both coolwater families such as Percidae and Catostomidae, and warmwater Centrarchidae (Kelso and Johnson 1991). Only a few species make up most of the biomass in Barrenlands lakes. A relatively small decline in the relative abundance of the dominant species with increasing species richness indicates an uneven distribution of species even in Barrenlands lakes with multiple species. In contrast, we expected the dominant species to contribute less to the total
biomass in Ontario lakes because the total number of species is higher; however, we found the contribution of the dominant species to total fish biomass to be very similar to Barrenlands lakes. It appears that only a few species contribute most of the biomass in both areas.

4.5.2.1 Lake trout

We conducted further analyses of lake trout because of its dominance in the Barrenlands. Although lake trout dominated only 34% of Barrenlands lakes in abundance, it dominated 58% of lakes by biomass. Two sets of information are relevant when predicting a latitudinal cline in lake trout biomass. An extension of the species-energy hypothesis predicts that abundance and biomass will decline with increasing latitude due to decreasing energy availability (Wright 1983; Srivastava and Lawton 1998). In general, coldwater lakes are oligotrophic and have low primary productivity, which limits fish abundance. Based on this information, lake trout biomass should be lower in the northern lakes. However, lake trout are coldwater fish and prefer summer water temperatures of 8 – 12 ºC (Christie and Regier 1988; Magnuson et al. 1990). Based solely on their thermal preference and the amount of habitat available to them at those temperatures, lake trout biomass should be higher in coldwater northern lakes than warmer southern lakes. Therefore, our finding of no significant latitudinal relationship in lake trout biomass suggests that these two forces may balance each other out.

Previous studies in southern Canada (i.e., Ontario) have shown that lake trout is more likely to occur in larger lakes (Ryan and Marshall 1994) and larger lakes contain larger lake trout (Shuter et al. 1998). These observations are consistent with our finding that larger lakes have higher lake trout biomass than smaller lakes (the slope of the regression between lake area and biomass was significantly higher than 1). Large lakes are typically deeper than small lakes and this increase in depth would typically add to the additional suitable habitat that is directly associated with
increased lake size. Also, in warm summer months, large lakes with greater depths thermally stratify and thus provide hypolimnetic coldwater habitat for lake trout during this season. In winter, deep lakes contain more habitats for lake trout because deep lakes contain a higher proportion of water (in liquid form) compared to small, shallow lakes which are more susceptible to freezing (Ryan and Marshall 1994).

4.5.3 Future of northern fishes

In summary, our findings suggest that northern Canadian lakes support higher fish diversity and resources than might be expected based on species-energy theory. Many studies predict that with climate change, water temperatures in southern Canadian lakes will exceed optimal temperatures for many fish species, resulting in local extirpations or range contractions (e.g., Schindler 2001; Reist et al. 2006a; Eliason et al. 2011). For example, declines of salmonids in southern temperate lakes are predicted due to warming waters (Clews et al. 2010; Pankhurst and King 2010; Blair et al. 2013). As a result, northern lakes represent an important refugia for salmonid fishes that represent significant commercial, recreational and aboriginal value, as well as biogeographical reservoirs of biodiversity. Furthermore, northern Canadian lakes, including Barrenlands lakes, are occupied by long-living, slow-growing fishes (Johnson 1976; Sparholt 1985), which are more vulnerable to habitat loss (or any other disturbances) than southern fishes because the recovery from a disturbance will likely require a longer period of time.

Fishes occupying Barrenland lakes continue to be threatened by ongoing mineral resource extraction projects in the Barrenland region (Mackenzie Valley Environmental Impact Review Board 2013). Recent changes to the Canadian Fisheries Act may increase the risk to northern fish populations (Hutchings and Post 2013), but this will depend greatly on how the new Act is enforced. The interpretation of the new Act expressed by Hutchings and Post (2013) suggests
Chapter 4: Fish diversity and biomass in northern Canadian lakes

that aquatic habitats without a “fishery” (e.g. the Barrenlands lakes) may lose their ‘right’ to protection in order to speed up mineral extraction projects in Northern Canada. Our findings highlight (1) the value of maintaining northern lakes for the future as enclaves of fish biodiversity in Canada, and (2) the need for new research to better understand northern lakes and the fishes they support in the face of ongoing development.

4.6 Acknowledgements

The authors would like to thank everyone who contributed to data collection, Shelly Boss for creation of the database for DFO, Don Jackson for sharing comparative data for analyses, Bruce Hanna for providing additional information, Helen Rodd and two anonymous reviewers for providing constructive feedback on the manuscript. Samaranin was supported by NSERC grants (to B. Shuter and H. Rodd) and by the Department of Ecology & Evolutionary Biology at the University of Toronto.

4.7 Literature cited


Chapter 4: Fish diversity and biomass in northern Canadian lakes


Chapter 4: Fish diversity and biomass in northern Canadian lakes


Chapter 4: Fish diversity and biomass in northern Canadian lakes


Chapter 4: Fish diversity and biomass in northern Canadian lakes


Chapter 4: Fish diversity and biomass in northern Canadian lakes


Chapter 4: Fish diversity and biomass in northern Canadian lakes


Chapter 4: Fish diversity and biomass in northern Canadian lakes


### 4.8 Table

**Table 4.1: Summary of main analyses and information on data used for the analysis**

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Data included in dataset</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Preliminary regional species richness of Barrenlands lakes</td>
<td>All 53 lakes/ponds in our study area</td>
<td>Barrenlands fish-out database, Fisheries &amp; Oceans Canada</td>
</tr>
<tr>
<td>b) Species richness estimation of Barrenlands monitored lakes</td>
<td>Monitored lakes with &gt; 95 individuals sampled</td>
<td>Barrenlands fish-out database, Fisheries &amp; Oceans Canada</td>
</tr>
<tr>
<td>e) Abundance and biomass estimation</td>
<td>Barrenlands lakes &gt;10 ha and &gt;50 fish sampled</td>
<td>Barrenlands fish-out database, Fisheries &amp; Oceans Canada</td>
</tr>
</tbody>
</table>
## Chapter 4: Fish diversity and biomass in northern Canadian lakes

<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>g) Overall biomass analysis</td>
<td>Barrenlands + other lake fish biomass data from literature</td>
<td>see Appendix 1 for individual sources. Barrenlands fish-out database; Kelso &amp; Johnson 1991; Downing &amp; Plante 1993; Randall et al. 1995; Achleitner et al. 2012</td>
</tr>
<tr>
<td>h) Latitudinal pattern in lake trout Biomass</td>
<td>Lake trout biomass data from Barrenlands and other Canadian lakes</td>
<td>Barrenlands fish-out database; Lester, N.P. Personal communication</td>
</tr>
</tbody>
</table>
Figure 4.1: Location of the study area in Northwest Territories, Canada. Northern Canada (north of 60°N latitude) contains the three Canadian territories: Nunavut, Northwest Territories, and Yukon.
Figure 4.2: Occurrence of each fish species in 53 lakes and ponds in the Barrenlands region of Northwest Territories, Canada. ARGR: Arctic grayling; BURB: Burbot; CISC: Cisco; LKCH: Lake chub; LKTR: Lake trout; LKWH: Lake whitefish; LNSC: Longnose sucker; RNWH: Round whitefish; SLSC: Slimy sculpin.
Figure 4.3: Relationship between lake area and species richness for Barrenlands lakes. Dark circles are species richness estimates for fish-out lakes and open circles are species richness estimates for monitored lakes. Regression line shows the overall trend for all lakes.
Figure 4.4: Relationship between species richness and lake area for lakes in different regions in North America. There was no statistical difference in slopes among regions but intercepts were significantly different. The common slope was plotted with the region specific intercept because slopes were not significantly different. Positioning of the symbol indicates the average lake size for that region. Note that regression lines for Black-Hollow region and Manitoulin Island coincide.
Figure 4.5: Total fish biomass distributions for Barrenlands lakes (light grey) and Ontario lakes (dark grey) on a shared axis. The two distributions were very marginally different ($p = 0.08$; Wilcoxon rank-sum test).
Figure 4.6: Relationship between fish biomass, lake area, mean annual air temperature, and species richness.  
a) Log$_{10}$ transformed total fish biomass as a function of Log$_{10}$ transformed lake area. Dark circles represent Barrenlands lakes.  
b) Log$_{10}$ transformed biomass per area (in kg/ha) as a function of mean annual air temperature with numerical symbols representing the species richness. Barrenlands lakes are located around -10 °C.
Figure 4.7: a) Mean percentage abundance of the dominant species for lakes with more than one species. b) Mean percentage biomass of the dominant species for lakes with more than one species. Dark circles indicate the mean percentage biomass of the dominant species from Ontario lakes (calculated from Kelso & Johnson 1991). Error bars represent the standard error of the mean.
Chapter 4: Fish diversity and biomass in northern Canadian lakes

Figure 4.8: Lake trout biomass in Canadian lakes. There was no a significant linear relationship between lake trout biomass and latitude ($R^2 = 0.01, p = 0.44$). Dark circles represent lake trout biomass in Barrenlands lakes.
## Chapter 4: Fish diversity and biomass in northern Canadian lakes

### 4.10 Appendix

A 4.1: Summary of data used for analysis of fish biomass

<table>
<thead>
<tr>
<th>Lake name</th>
<th>Country</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Lake Area (ha)</th>
<th>Total Biomass (kg)</th>
<th>Mean annual air temperature (ºC)</th>
<th>Fish Species Richness</th>
<th>Mean depth (m)</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marion</td>
<td>Canada (BC)</td>
<td>50.99°</td>
<td>-127.00°</td>
<td>13</td>
<td>375.7</td>
<td>7.91</td>
<td>2</td>
<td>2.4</td>
<td>Efford 1972</td>
</tr>
<tr>
<td>Airstrip*</td>
<td>Canada (NT)</td>
<td>64.68°</td>
<td>-110.59°</td>
<td>19.3</td>
<td>306.3</td>
<td>-10.24</td>
<td>3</td>
<td>1.5</td>
<td>This paper</td>
</tr>
<tr>
<td>Beartooth†</td>
<td>Canada (NT)</td>
<td>64.73°</td>
<td>-110.58°</td>
<td>4.8</td>
<td>36.3</td>
<td>-10.24</td>
<td>2</td>
<td>3.25</td>
<td>This paper</td>
</tr>
<tr>
<td>Big Pond†</td>
<td>Canada (NT)</td>
<td>64.81°</td>
<td>-110.61°</td>
<td>28.2</td>
<td>23.7</td>
<td>-10.24</td>
<td>1</td>
<td>0.8</td>
<td>This paper</td>
</tr>
<tr>
<td>Brandy*</td>
<td>Canada (NT)</td>
<td>64.72°</td>
<td>-110.74°</td>
<td>29.9</td>
<td>240.8</td>
<td>-10.24</td>
<td>5</td>
<td>ND</td>
<td>This paper</td>
</tr>
<tr>
<td>Counts†</td>
<td>Canada (NT)</td>
<td>64.65°</td>
<td>-110.29°</td>
<td>77.5</td>
<td>1270.4</td>
<td>-10.36</td>
<td>2</td>
<td>1.9</td>
<td>This paper</td>
</tr>
<tr>
<td>Cujo†</td>
<td>Canada (NT)</td>
<td>64.58°</td>
<td>-110.19°</td>
<td>44.3</td>
<td>591.8</td>
<td>-10.36</td>
<td>3</td>
<td>2.2</td>
<td>This paper</td>
</tr>
<tr>
<td>Fay Bay†</td>
<td>Canada (NT)</td>
<td>64.75°</td>
<td>-110.69°</td>
<td>46.4</td>
<td>462.2</td>
<td>-10.24</td>
<td>3</td>
<td>2.47</td>
<td>This paper</td>
</tr>
<tr>
<td>Grizzly†</td>
<td>Canada (NT)</td>
<td>64.72°</td>
<td>-110.55°</td>
<td>58.9</td>
<td>361.9</td>
<td>-10.24</td>
<td>2</td>
<td>14.6</td>
<td>This paper</td>
</tr>
<tr>
<td>Horseshoe†</td>
<td>Canada (NT)</td>
<td>64.87°</td>
<td>-110.54°</td>
<td>75.7</td>
<td>521.7</td>
<td>-10.24</td>
<td>3</td>
<td>2</td>
<td>This paper</td>
</tr>
<tr>
<td>Koala*</td>
<td>Canada (NT)</td>
<td>64.72°</td>
<td>-110.60°</td>
<td>38.0</td>
<td>372.3</td>
<td>-10.24</td>
<td>5</td>
<td>5.9</td>
<td>This paper</td>
</tr>
<tr>
<td>Kodiak†</td>
<td>Canada (NT)</td>
<td>64.70°</td>
<td>-110.62°</td>
<td>90.7</td>
<td>995.9</td>
<td>-10.24</td>
<td>4</td>
<td>2.1</td>
<td>This paper</td>
</tr>
<tr>
<td>Logan†</td>
<td>Canada (NT)</td>
<td>64.87°</td>
<td>-110.77°</td>
<td>127.2</td>
<td>1204.1</td>
<td>-10.24</td>
<td>2</td>
<td>5.81</td>
<td>This paper</td>
</tr>
<tr>
<td>Long*</td>
<td>Canada (NT)</td>
<td>64.70°</td>
<td>-110.70°</td>
<td>614.4</td>
<td>8411.1</td>
<td>-10.24</td>
<td>8</td>
<td>7.4</td>
<td>This paper</td>
</tr>
</tbody>
</table>
Chapter 4: Fish diversity and biomass in northern Canadian lakes

<table>
<thead>
<tr>
<th>Location</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Diversity</th>
<th>Biomass</th>
<th>Depth</th>
<th>Year</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misery*</td>
<td>64.56°</td>
<td>-110.18°</td>
<td>13.7</td>
<td>22.1</td>
<td>-10.24</td>
<td>1</td>
<td>This paper</td>
</tr>
<tr>
<td>Moose†</td>
<td>64.68°</td>
<td>-110.64°</td>
<td>43.6</td>
<td>569.4</td>
<td>-10.24</td>
<td>4</td>
<td>ND This paper</td>
</tr>
<tr>
<td>Nanuq†</td>
<td>64.91°</td>
<td>-110.29°</td>
<td>189.5</td>
<td>2188.7</td>
<td>-10.36</td>
<td>2</td>
<td>ND This paper</td>
</tr>
<tr>
<td>Nema†</td>
<td>64.67°</td>
<td>-110.71°</td>
<td>77.5</td>
<td>1330.1</td>
<td>-10.24</td>
<td>5</td>
<td>1.9 This paper</td>
</tr>
<tr>
<td>Panda*</td>
<td>64.72°</td>
<td>-110.58°</td>
<td>35.0</td>
<td>90.3</td>
<td>-10.36</td>
<td>4</td>
<td>3.8 This paper</td>
</tr>
<tr>
<td>Rene†</td>
<td>64.65°</td>
<td>-110.81°</td>
<td>32.9</td>
<td>332.1</td>
<td>-10.24</td>
<td>4</td>
<td>ND This paper</td>
</tr>
<tr>
<td>Sable†</td>
<td>64.86°</td>
<td>-110.51°</td>
<td>8.9</td>
<td>39.1</td>
<td>-10.24</td>
<td>1</td>
<td>6.3 This paper</td>
</tr>
<tr>
<td>Slipper†</td>
<td>64.62°</td>
<td>-110.86°</td>
<td>189.5</td>
<td>3802.8</td>
<td>-10.24</td>
<td>4</td>
<td>ND This paper</td>
</tr>
<tr>
<td>Two Rock†</td>
<td>64.86°</td>
<td>-110.53°</td>
<td>28.6</td>
<td>404.2</td>
<td>-10.24</td>
<td>2</td>
<td>3.6 This paper</td>
</tr>
<tr>
<td>Ulu†</td>
<td>64.86°</td>
<td>-110.51°</td>
<td>28.7</td>
<td>415.2</td>
<td>-10.24</td>
<td>3</td>
<td>2.7 This paper</td>
</tr>
<tr>
<td>Vulture†</td>
<td>64.75°</td>
<td>-110.54°</td>
<td>180.2</td>
<td>4512.8</td>
<td>-10.24</td>
<td>2</td>
<td>ND This paper</td>
</tr>
<tr>
<td>White†</td>
<td>64.70°</td>
<td>-110.76°</td>
<td>57.2</td>
<td>985.8</td>
<td>-10.24</td>
<td>3</td>
<td>1.9 This paper</td>
</tr>
<tr>
<td>Willy*</td>
<td>64.72°</td>
<td>-110.72°</td>
<td>23.9</td>
<td>248.2</td>
<td>-10.24</td>
<td>4</td>
<td>ND This paper</td>
</tr>
<tr>
<td>Char</td>
<td>74.71°</td>
<td>-94.90°</td>
<td>53</td>
<td>487.6</td>
<td>-16.83</td>
<td>1</td>
<td>10.2 MacCallum &amp; Regier 1984</td>
</tr>
<tr>
<td>04 Pender</td>
<td>45.32°</td>
<td>-79.87°</td>
<td>22.7</td>
<td>442.65</td>
<td>5.44</td>
<td>4</td>
<td>5.3 Kelso &amp; Johnson 1991</td>
</tr>
<tr>
<td>14NO Carruthers</td>
<td>45.37°</td>
<td>-79.87°</td>
<td>10</td>
<td>1199</td>
<td>5.44</td>
<td>8</td>
<td>2.7 Kelso &amp; Johnson 1991</td>
</tr>
<tr>
<td>20 Cochrane</td>
<td>45.32°</td>
<td>-79.94°</td>
<td>18.9</td>
<td>625.59</td>
<td>5.44</td>
<td>7</td>
<td>3.2 Kelso &amp; Johnson 1991</td>
</tr>
<tr>
<td>33 Lane</td>
<td>45.35°</td>
<td>-79.93°</td>
<td>18.5</td>
<td>247.9</td>
<td>5.44</td>
<td>5</td>
<td>3.1 Kelso &amp; Johnson 1991</td>
</tr>
<tr>
<td>A1 Orange</td>
<td>45.60°</td>
<td>-79.90°</td>
<td>9.8</td>
<td>292.04</td>
<td>4.53</td>
<td>4</td>
<td>2.1 Kelso &amp; Johnson 1991</td>
</tr>
<tr>
<td>Lake Name</td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Depth (m)</td>
<td>Water (°C)</td>
<td>Oxygen (mg/L)</td>
<td>Dominant Fish</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------</td>
<td>--------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Big Turkey</td>
<td>Canada (ON)</td>
<td>47.05°</td>
<td>-84.42°</td>
<td>52</td>
<td>379.6</td>
<td>2.94</td>
<td></td>
</tr>
<tr>
<td>C1 Crow</td>
<td>Canada (ON)</td>
<td>45.78°</td>
<td>-80.23°</td>
<td>11.5</td>
<td>223.1</td>
<td>5.16</td>
<td></td>
</tr>
<tr>
<td>K12 Raven</td>
<td>Canada (ON)</td>
<td>45.53°</td>
<td>-80.20°</td>
<td>13.6</td>
<td>137.36</td>
<td>5.16</td>
<td></td>
</tr>
<tr>
<td>K25 Turbid</td>
<td>Canada (ON)</td>
<td>46.12°</td>
<td>-81.19°</td>
<td>17</td>
<td>323</td>
<td>4.78</td>
<td></td>
</tr>
<tr>
<td>Little Turkey</td>
<td>Canada (ON)</td>
<td>47.04°</td>
<td>-84.41°</td>
<td>19</td>
<td>212.8</td>
<td>2.94</td>
<td></td>
</tr>
<tr>
<td>Lonesome</td>
<td>Canada (ON)</td>
<td>45.35°</td>
<td>-75.05°</td>
<td>3.6</td>
<td>182.88</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>Long Pond</td>
<td>Canada (ON)</td>
<td>42.58°</td>
<td>-80.42°</td>
<td>2.4</td>
<td>666.72</td>
<td>7.98</td>
<td></td>
</tr>
<tr>
<td>Pine</td>
<td>Canada (ON)</td>
<td>45.35°</td>
<td>-75.05°</td>
<td>16.4</td>
<td>577.28</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>Red deer</td>
<td>Canada (ON)</td>
<td>45.53°</td>
<td>-79.14°</td>
<td>6</td>
<td>69.6</td>
<td>4.10</td>
<td></td>
</tr>
<tr>
<td>Redpine</td>
<td>Canada (ON)</td>
<td>45.35°</td>
<td>-75.05°</td>
<td>8.3</td>
<td>420.81</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>Sawlog</td>
<td>Canada (ON)</td>
<td>45.35°</td>
<td>-75.05°</td>
<td>6.3</td>
<td>310.59</td>
<td>5.50</td>
<td></td>
</tr>
<tr>
<td>Wishart</td>
<td>Canada (ON)</td>
<td>47.04°</td>
<td>-84.40°</td>
<td>19</td>
<td>463.6</td>
<td>2.94</td>
<td></td>
</tr>
<tr>
<td>Achensee</td>
<td>Austria</td>
<td>47.46°</td>
<td>11.71°</td>
<td>680</td>
<td>22984</td>
<td>4.38</td>
<td></td>
</tr>
<tr>
<td>Altausseer</td>
<td>Austria</td>
<td>47.64°</td>
<td>13.78°</td>
<td>210</td>
<td>7098</td>
<td>5.68</td>
<td></td>
</tr>
<tr>
<td>Attersee</td>
<td>Austria</td>
<td>47.88°</td>
<td>13.55°</td>
<td>4620</td>
<td>216678</td>
<td>5.68</td>
<td></td>
</tr>
<tr>
<td>Faaker See</td>
<td>Austria</td>
<td>46.58°</td>
<td>13.93°</td>
<td>220</td>
<td>17226</td>
<td>4.43</td>
<td></td>
</tr>
<tr>
<td>Fuschlsee</td>
<td>Austria</td>
<td>47.80°</td>
<td>13.28°</td>
<td>265</td>
<td>56445</td>
<td>5.97</td>
<td></td>
</tr>
<tr>
<td>Grundlsee</td>
<td>Austria</td>
<td>47.63°</td>
<td>13.86°</td>
<td>410</td>
<td>20459</td>
<td>5.68</td>
<td></td>
</tr>
</tbody>
</table>

Chapter 4: Fish diversity and biomass in northern Canadian lakes

<table>
<thead>
<tr>
<th>Location</th>
<th>Country</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
<th>Biomass (t)</th>
<th>Diversity</th>
<th>Total N</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Millstatter</td>
<td>Austria</td>
<td>46.80°</td>
<td>13.57°</td>
<td>1328</td>
<td>113411.2</td>
<td>4.43</td>
<td>ND</td>
<td>Achleitner et al 2012</td>
</tr>
<tr>
<td>Obertrumer</td>
<td>Austria</td>
<td>47.96°</td>
<td>13.09°</td>
<td>488</td>
<td>34355.2</td>
<td>5.97</td>
<td>ND</td>
<td>Achleitner et al 2012</td>
</tr>
<tr>
<td>Toplitzsee</td>
<td>Austria</td>
<td>47.64°</td>
<td>13.93°</td>
<td>54</td>
<td>27</td>
<td>5.68</td>
<td>ND</td>
<td>Achleitner et al 2012</td>
</tr>
<tr>
<td>Wolfgangsee</td>
<td>Austria</td>
<td>47.73°</td>
<td>13.44°</td>
<td>1284</td>
<td>69464.4</td>
<td>5.97</td>
<td>ND</td>
<td>Achleitner et al 2012</td>
</tr>
<tr>
<td>Worthersee</td>
<td>Austria</td>
<td>46.62°</td>
<td>14.16°</td>
<td>1939</td>
<td>50026.2</td>
<td>7.13</td>
<td>ND</td>
<td>Achleitner et al 2012</td>
</tr>
<tr>
<td>La Luisa</td>
<td>Cuba</td>
<td>22.72°</td>
<td>-82.13°</td>
<td>1</td>
<td>325.7</td>
<td>24.37</td>
<td>3</td>
<td>1.4 Hoiafk 1970</td>
</tr>
<tr>
<td>Alinen</td>
<td>Austria</td>
<td>46.80°</td>
<td>13.57°</td>
<td>1328</td>
<td>113411.2</td>
<td>4.43</td>
<td>ND</td>
<td>Achleitner et al 2012</td>
</tr>
<tr>
<td>Mustajärvi</td>
<td>Finland</td>
<td>61.56°</td>
<td>22.04°</td>
<td>1</td>
<td>73</td>
<td>4.03</td>
<td>3</td>
<td>Rask and Arvola 1985</td>
</tr>
<tr>
<td>Horkkajarvi</td>
<td>Finland</td>
<td>61.21°</td>
<td>25.16°</td>
<td>1</td>
<td>32.2</td>
<td>3.60</td>
<td>2</td>
<td>Rask and Arvola 1985; Arvola 1983</td>
</tr>
<tr>
<td>Iso Mustajärvi</td>
<td>Finland</td>
<td>62.12°</td>
<td>23.47°</td>
<td>3</td>
<td>58.5</td>
<td>3.69</td>
<td>4</td>
<td>Lappalainen et al 1988</td>
</tr>
<tr>
<td>Nakuru</td>
<td>Kenya</td>
<td>-0.36°</td>
<td>36.10°</td>
<td>4400</td>
<td>261360</td>
<td>15.25</td>
<td>ND</td>
<td>Vareschi and Jacobs 1984, 1985; Plante 1987</td>
</tr>
<tr>
<td>Ovre Heimdalsvatn</td>
<td>Norway</td>
<td>61.42°</td>
<td>8.87°</td>
<td>78</td>
<td>1544.4</td>
<td>-0.92</td>
<td>2</td>
<td>Kloster 1987; Lien 1978, 1981; Tangen and Bretum 1978</td>
</tr>
<tr>
<td>Dalnee</td>
<td>Russia</td>
<td>47.50°</td>
<td>47.06°</td>
<td>136</td>
<td>26792</td>
<td>-1.12</td>
<td>3</td>
<td>Krogius et al. 1972</td>
</tr>
<tr>
<td>Demenets</td>
<td>Russia</td>
<td>56.64°</td>
<td>29.69°</td>
<td>7</td>
<td>2605.4</td>
<td>4.55</td>
<td>7</td>
<td>Gulin and Rudenko 1973</td>
</tr>
<tr>
<td>Botjarn</td>
<td>SWE</td>
<td>60.27°</td>
<td>15.07°</td>
<td>10</td>
<td>303</td>
<td>3.14</td>
<td>1</td>
<td>Nyberg 1979; Ramberg 1976</td>
</tr>
<tr>
<td>Coleback</td>
<td>USA</td>
<td>45.02°</td>
<td>-67.21°</td>
<td>9</td>
<td>172.8</td>
<td>5.42</td>
<td>5</td>
<td>Rupp &amp; DeRoche 1965</td>
</tr>
<tr>
<td>Hills</td>
<td>USA</td>
<td>44.56°</td>
<td>-70.18°</td>
<td>8</td>
<td>58.4</td>
<td>ND</td>
<td>5</td>
<td>Rupp &amp; DeRoche 1965</td>
</tr>
<tr>
<td>Shagg</td>
<td>USA</td>
<td>44.42°</td>
<td>-70.53°</td>
<td>23</td>
<td>167.9</td>
<td>4.54</td>
<td>6</td>
<td>Rupp &amp; DeRoche 1965</td>
</tr>
</tbody>
</table>
Chapter 4: Fish diversity and biomass in northern Canadian lakes

| Washington  | USA | 47.61° | -122.26° | 8760 | 818184 | 11.21 | 8 | 33 | Edmondson 1977, Eggers et al. 1978 |

*: Fishout lake from Barrenlands

†: Monitored lake from Barrenlands

ND: No data
4.11 Copyright Acknowledgement

This chapter was published as: Samarasin P., Minns CK., Shuter BJ., Tonn WM., and Rennie MD., 2015. Fish species diversity and biomass in northern and southern Canadian lakes: northern lakes are more diverse and have greater biomass than expected based on species-energy theory. Canadian Journal of Fisheries & Aquatic Sciences. 72(2): 226-237.
Chapter 5: Conclusions

Chapter 5

Conclusions

Habitat loss, alteration and degradation are considered to be the most important threats affecting fish and wildlife today (Wilcove et al. 1998; Balmford et al. 2005). Although the effects of habitat loss on genetic, species, and community diversity have been documented in empirical studies, mechanisms of how these changes lead to biodiversity loss and procedures for mitigating these negative consequences are still under investigation by conservation biologists worldwide (Tilman et al. 1994; Brook et al. 2003; Wake & Vredenburg 2008). In this thesis, I examined the effects of recent changes in habitat on all three levels of diversity. These effects are a particularly vital area of research because of unprecedented, global habitat changes within the last century and the effects such changes can have on fish and wildlife populations and communities.

I began my thesis research with a very basic question: what are the effects of recent changes in population connectivity, caused by habitat change, on our ability to accurately infer current levels of population connectivity? I was curious to know how genetic methods that estimate connectivity, which are increasingly used to infer connectivity, perform in situations where connectivity has changed in the recent past. To answer this question, I used a simulation approach to evaluate the effects of recent changes in population connectivity on the accuracy of methods commonly used to infer genetic connectivity (Chapter 2). As far as I am aware, my study is the first to evaluate this potential problem. My results indicate that coalescent estimates should not be interpreted as estimates of historical, long-term connectivity as they tend to underestimate historical connectivity when there has been a recent decline in connectivity. In contrast, disequilibrium-based methods (e.g., BayesAss), which are usually used to estimate
Chapter 5: Conclusions

contemporary connectivity, are likely to overestimate current connectivity in such situations. Therefore, my results indicated that researchers need to be cautious about inferring connectivity from these methods when there is reason to believe that connectivity has changed in the recent past. In particular, inferring changes in connectivity that have occurred in the recent past by contrasting estimates from coalescent and disequilibrium methods can lead to misleading conclusions.

Shifting from simulated to empirical population data, I continued my research on connectivity among populations in changing habitats. Here, I investigated population genetic and life-history effects of hydroelectric dams on Sockeye salmon populations in the lower Fraser River, B.C. (Chapter 3). This case study provided important insights into these issues because, i) construction of hydroelectric dams in early the 1900’s is thought to have changed connectivity among populations, ii) historical records indicated that, because of dam construction, anadromous Sockeye populations had either been lost or possibly converted to the resident form (i.e., kokanee), and iii) recent observations showed that some presumed “kokanee” were migrating to the ocean during experimental water releases. These circumstances made this case interesting from ecological, evolutionary and conservation perspectives and an excellent example of a recent change in connectivity resulting from habitat change. My genetic analyses of the connectivity of 11 putative Sockeye populations indicated first that, for many populations, including ones that are not in the immediate vicinity of the dams, there have been changes in connectivity between historical and contemporary time periods. This suggests that even the connectivity of populations not in the immediate vicinity of hydroelectric dams can be altered indirectly because of changes in water flow or chemical cues used by salmon for natal migration. In addition, my analyses indicated a genetic distinction between sea-run and resident individuals
in the Coquitlam reservoir, raising further intriguing research questions for the future. It appears that some descendants of the anadromous Sockeye salmon that were trapped by the dam did not lose their ability to migrate to the ocean over a period of almost 100 years. But my results also indicate very early stages of genetic differentiation between sea-run and resident individuals within the reservoir, meaning that some could be converting to the resident form. These results also suggest that there could be a genetic basis for anadromous and resident life history types; which continues to be investigated with limited success in salmonid species (Perrier et al. 2013; Sutherland et al. 2014).

In Chapter 4, moving from a single species to a community focus, I investigated fish species diversity and biomass in northern Canadian lakes. This is one of only few studies to quantify fish diversity and biomass in the subarctic region. I was able to take advantage of monitoring and exhaustive sampling data that were collected for a proposed diamond mine development in the area. The collected data had been used only sparingly for an environmental impact assessment and did not appear to have been analyzed in much detail. Surprisingly, I found that these lakes contained higher species diversity than expected, given the latitude. Also surprisingly, I found that the species-area relationships for northern and southern Canadian lakes had similar slopes; in other words, contrary to expectation, I found that the rate at which species richness increased with lake area was similar for northern and southern Canadian lakes. I also found similar fish biomass distributions in northern and southern Canadian lakes. Overall, these results indicate that the fish diversity and resources in Northern Canada are higher than expected based on species-energy theory. Therefore, these results emphasize the value of conserving such lakes, especially in light of climate change, which is expected to affect many cold water populations living near the southern edge of their thermal tolerance.
5.1 Future directions

5.1.1 Understanding effects of recent connectivity changes on genetic inferences

Understanding how recent changes to population connectivity can influence genetic inferences of connectivity is important for biodiversity conservation because molecular methods are being used with increasing frequency. We need a better understanding of the limitations of these methods to prevent incorrect inferences and conclusions about the status of threatened and endangered species; incorrect information could have serious consequences for risk assessment and recovery of wildlife populations. As more computational power becomes available in the future, simulation approaches can be used to investigate other biases resulting from recent changes in connectivity on genetic inferential methods. Other analytical methods used by molecular ecologists and conservation biologists should be tested for their performance under similar conditions. By testing these methods, we can identify best methods to accurately estimate connectivity when connectivity has changed in recent times. Building on my research, other scenarios could be simulated; for example, what are the effects of a recent increase in gene flow that can result when previously isolated populations come into contact with each other as habitat becomes limiting? This is an important question because there are conservation concerns about the effects of interbreeding of historically diverged populations; for example loss of fitness through outbreeding depression and loss of recognized species through introgression. In this kind of analysis, it would be useful to investigate the effects of the number of loci used and the number of individuals sampled on estimate accuracy. Such studies would be of broader interest because population genomic studies sample very few individuals per population but obtain large segments of those individual’s genomes. Consequently, knowledge of the relative importance of loci versus individuals would be of practical importance. Like migration rate estimators, genetic
clustering methods could also be affected by the timing of connectivity changes. Therefore, it
would be interesting to investigate how the timing of connectivity changes affect clustering
methods such as STRUCTURE and DAPC, which are used to identify the population genetic
structure of species, and determine if the accuracy and bias trends are consistent among
migration rate estimators and genetic clustering methods.

5.1.2 Evolution of life-history types in Sockeye salmon

There are multiple avenues of research that can be explored further to understand the genetics of
anadromous and kokanee life-history types in Sockeye salmon. Building on my research, the
microsatellite loci that I identified as outliers and that differentiated between sea-run and resident
individuals can be used as a starting point to conduct genetic linkage mapping and to explore
genes near those microsatellite loci. This may lead to identification of genes, or genetic regions,
that differ between the two life history types. Another approach would be to perform RAD
sequencing with a few “sea-run” and “resident” individuals that were assigned to each category
with the highest assignment probability. This might identify SNPs that are fixed within a group
but that differ between the two life history types. Also, the recently sequenced rainbow trout
(Oncorhynchus mykiss) genome (Berthelot et al. 2014) will provide more tools for understanding
genomic architecture differences behind the two life history types of Sockeye salmon and will
also provide a reference genome for assembling the Sockeye genome when it is sequenced.

5.1.3 Quantifying aquatic resources in the northern Canada

My study provided some insights into fish diversity and resources in subarctic Canada and how
they compare with southern Canada, however, more studies from across Northern Canada are
needed. Additional data from different parts of the three northern territories would be helpful for
elucidating the relative importance of biogeographical and local factors to the diversity of lake
Chapter 5: Conclusions

fishes. Our results, at a restricted spatial scale, suggest that regional factors (e.g. biogeography) are more important in explaining diversity than local factors (e.g. lake area). A multivariate study could help determine the relative importance of climatic, biogeographic, and local factors in determining the fish diversity in Northern Canada.

The data used in my study were collected in 1990’s to early 2000’s and, except to identify what species were present in the area, the data had not been analyzed in any detail. I expect that similar data were collected by consulting companies for the resource extraction industry and have not being analyzed. It might be possible to use these sampling data to estimate abundance, biomass, and possibly productivity in northern lakes at the population, species, and community levels, depending on the availability of data. For example, fish removal data from Kennady Lake for the Gahcho Kué diamond mine project would likely produce types of data that would be similar to ours and would enhance documentation of northern fish communities. Other development proposals and subsequent monitoring projects in the north should be required to provide biodiversity data that can be used to get a comprehensive understanding of diversity and wildlife resources in Northern Canada.

5.2 Final remarks

As a conservation biologist, my goal for this thesis was to address gaps in our knowledge that would contribute to conserving biodiversity. A better understanding of the ecology and evolution of target taxa and of conservation tools is needed to achieve our conservation objectives. Therefore, using simulation and empirical approaches, I focused my investigations on several conservation-related issues that are linked to the ecology and evolution of animals. I addressed conservation issues at different scales, from populations to communities, because improving our knowledge at all levels is important for conservation. Most importantly, drastic changes in the
Chapter 5: Conclusions

quantity and quality of habitat continue to be a threat to the persistence of many fish and wildlife populations, species, and communities today. Habitats are affected in a variety of ways, from direct habitat destruction, to alteration and degradation through human development, pollution and climate change. Consequently, my thesis mainly focused on investigations of the effects of habitat change on biodiversity in space and in time, with a significant emphasis on how habitat change can influence populations through a loss of connectivity. The aims and motivations behind my thesis are well captured by the following quote by Dr. E. O. Wilson:

“The one process now going on that will take millions of years to correct is the loss of genetic and species diversity by the destruction of natural habitats. This is the folly our descendants are least likely to forgive us.”

5.3 Literature cited


