Genetic and phenotypic variation of Atlantic salmon (*Salmo salar*) young-of-the-year in the Upper Salmon River, NB, Canada

by

Emilie Diesbourg

A THESIS PREPARED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Science with Honours in Marine Biology

Supervisors: Kurt Samways, PhD, Department of Biological Sciences
and Canadian Rivers Institute
Cassidy D'Aloia, PhD, Department of Biology, University
of Toronto Mississauga

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit: http://creativecommons.org/licenses/by/4.0/ or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

THE UNIVERSITY OF NEW BRUNSWICK

SAINT JOHN

April, 2022

©Emilie Diesbourg, 2022

ABSTRACT

Determining the spatial scale at which genetic structuring occurs is one of the main goals of conservation genetics because such data can help identify management units. This study assessed the hierarchical genetic structure of Atlantic salmon young-of-the-year in the Upper Salmon River, NB at a fine spatial scale (~6 km). Using seven microsatellite loci, genetic clustering analyses and F_{ST} estimates were computed to determine the strength of genetic structure within the river. Patterns of genetic structure were then tested to detect whether they aligned with previously identified morphological groupings, using the same individuals. Subtle genetic differentiation (mean pairwise F_{ST}=0.036) and a weak isolation by distance pattern was found to exist within this ~6 km stretch of river. Additionally, not all morphological groupings were genetically distinct. This study highlights the importance of considering within-river genetic structure across small spatial scales for more informed conservation design.

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Kurt Samways and Cassidy D'Aloia for providing me with constant support and guidance throughout this project. I have learned so much from them both and have had so much fun with this project. They gave me constant support and motivation and I would not have been able to complete this study without their feedback. I would also like to acknowledge the work of Ashlee Prevost and Dr. Dylan Fraser, as well as Laura Clarke for providing me access to their datasets to complete my study.

STATEMENT OF RESEARCH CONTRIBUTION

Data collection for this project was done in the summer of 2017 by staff at Fundy National Park. Beginning in May, 2021 I planned out my research question and started the literature review for my project. By the end of the summer, September 2021, I compiled a library of literature for my project and organized multiple existing datasets on morphometrics, genetics, and environmental data, provided by Fundy National Park, into modified, usable datasets. Original datasets came from two independent studies by Prevost and Fraser, 2018, and Clarke, 2018. All data analysis from this study was done by myself, Emilie Diesbourg, and the statistical analyses were finished by the end of December 2021. All writing was completed by myself with edits from both of my supervisors. A full thesis draft was completed by the end of March 2022 and the final thesis was submitted in April 2022.

Table of Contents

List of Tables

List of Figures

Introduction

Phenotypic variation is the differential expression of phenotypes across individuals or populations and is ubiquitous among organisms (Hallgrimsson and Hall, 2005). Phenotypic variation is controlled genetically, environmentally, or through a combination of the two (Hallgrimsson and Hall, 2005). Because this variation provides a basis for evolution, understanding the genetic and environmental drivers of phenotypic variation and its evolutionary consequences is widely studied in evolutionary biology (Fusco and Minelli, 2010; Bull, 1987; Álvarez-Varas et al., 2021). A major goal in evolutionary biology is to determine whether observed phenotypic differences among individuals are driven by phenotypic plasticity or fixed genetic differences (Bull, 1987).

Phenotypic plasticity and genetic divergence are processes that coincide and interact with one another, potentially leading to trait divergence (Schmid and Guillaume, 2017). Phenotypic plasticity occurs when an individual's genotype expresses multiple phenotypes, induced by differing environmental conditions (Bradshaw, 1965; Price et al., 2003). Genetic divergence is a measure of how genetically distinct populations are and can eventually lead to intrinsic barriers to reproduction once there are enough genetic differences between lineages (Palumbi, 1994; Willem and Ferguson, 2002). Drivers of phenotypic plasticity include heterogeneity in environmental conditions and fluctuations in habitat quality, whereas drivers of genetic divergence include restricted historical and contemporary gene flow due to limited dispersal or strong natal homing (the tendency for individuals to return to their birthplace to reproduce; Stabell, 1984; Hansen et al., 1993). Genetic divergence is also more likely to arise when there are costs associated with a plastic response (Lucek et al., 2014; Fraser et al., 2011).

A species that offers a unique opportunity to explore the patterns, causes, and consequences of phenotypic variation among populations in both marine and freshwater habitats is Atlantic salmon (*Salmo salar*). Atlantic salmon is a species that holds high social, economic, and ecological importance, especially in North American and European cultures (Asche et al., 2005; Beland and Bielak, 2002). Despite their importance, Atlantic salmon populations have been precipitously declining for decades due to anthropogenic and environmental pressures, with the inner Bay of Fundy (iBoF) population of eastern Canada listed as endangered and protected under the Species At Risk Act (SARA) (DFO, 2010). Conservation of this species is challenging since many populations of Atlantic salmon experience variable environmental conditions and exhibit strong genetic divergence which may affect the strategy of preserving genetic diversity (Taylor, 1991). For example, when mitochondrial and microsatellite DNA are analyzed, a pattern of strong genetic structure is seen between the inner and outer Bay of Fundy populations on the east coast of Canada, and, therefore, these populations are treated differently when assessing their recovery potential (King et al., 2001). Populations of Atlantic salmon can also exhibit divergence of morphological traits, migratory behaviour, and reproductive ability, driven by local adaptation, which may have implications for management plans (Vincent et al., 2013).

Genetic divergence of Atlantic salmon populations has been documented at a large spatial scale between the North American and European continents (>1000 km; Davidson et al., 1989), as well as on smaller spatial scales, within river systems (100-1000 km; Dionne et al., 2008), and even at the tributary level (<100 km; Garant et al., 2000; Vähä et al., 2007; Kitanishi et al., 2009). Most research in this field focuses on genetic divergence between river systems rather than within river systems even though some studies indicate that gene flow can be restricted at fine spatial scales. For example, Jensen et al. (2017) found that Atlantic salmon populations within the Vefsna watershed (approximately 160km long) in Norway, exhibited evolution over less than 15 generations, while the only reproductively isolating barrier was natal homing. Additionally, evidence of genetic differentiation over a spatial scale less than 30 kilometers has been documented in other Salmonids, such as graylings (Haugen and Vollestad, 2000), brown trout (Carlsson et al., 1999), and masu salmon (Kitanishi et al., 2009), and may also apply to Atlantic salmon populations. Due to strong natal homing and the complex life history of the species, there is a need to further explore the fine-scale genetic structure of Atlantic salmon and determine its driver for proper management of the species.

Recently, wild origin Atlantic salmon adults and their offspring from the small Upper Salmon River (~a 12 km accessible river system), in Fundy National Park, New Brunswick, were genotyped for a parentage analysis (Prevost and Fraser, 2018). The morphological variation of the young-of-the-year (YOY) salmon was also quantified using geometric morphometrics (Clarke, 2018). This geometric study found that the YOY salmon in the Upper Salmon River displayed four distinct morphological groupings, some of which correspond to different regions of the river (Clarke, 2018). However, it is unknown whether the observed phenotypic variation between regions of the river is due to fine-scale local adaptation or due to phenotypic plasticity driven by differing environmental factors. Determining the driver of phenotypic variation is important to understand the evolutionary history of species, genetic/ecological processes that may lead to divergence and/or predicting population resilience to changing environmental pressures for conservation plans (Oostra et al., 2018; Gentili et al., 2018; Trzcinski et al., 2004).

This study examines whether the observed phenotypic variation among YOY Atlantic salmon in the Upper Salmon River is associated with genetic similarity between individuals. Firstly, this will be done by conducting a fine-scale genetic study whereby genetic groups are determined along the river, and secondly, by correlating the genotypes with the previously identified morphological groupings, using allele frequencies at seven microsatellite markers. It is predicted that, if multiple genetic groups exist along the river and phenotypic variation has a genetic basis, fish within a morphological grouping will be more genetically similar to each other than to fish in another morphological grouping. This study provides insight into fine-scale genetic structure within the population of Atlantic salmon in the Upper Salmon River, and the extent to which observed morphological variation correlates with genetic variation. The results from this study may be used as a future framework to determine the spatial scale at which YOY salmon from the Upper Salmon River exhibit local adaptation and provide feedback and management recommendations for the ongoing Fundy Salmon Recovery program.

Methods

Origin of data

Data for this study were obtained from two independent studies, Prevost and Fraser, (2018), and Clarke, (2018) which focused on genetic and morphological data, respectively, of YOY Atlantic salmon in the Upper Salmon River, Fundy National Park, New Brunswick, Canada. The YOY salmon were sampled in the summer of 2017 by electrofishing surveys at 20 locations spanning ~6 km of the ~12 km of accessible habitat of the Upper Salmon River. Tissue samples were collected from either the fish's adipose or caudal fin for genotyping (Prevost and Fraser, 2018). In total, 800 adults and 311 offspring were genotyped at seven polymorphic microsatellite loci: *SsaD144, SsaD157, SsaD48, SsaD71, SSsp2215, SSsp2216, SSspG7*, and a genetic parentage analysis was conducted to determine the parent-offspring relationships (Prevost and Fraser, 2018).

Prior to YOY individuals being sampled for genetic analysis, pictures were taken, and a geometric morphometric analysis was conducted using 16 different biological landmarks to assess phenotypic variation (Clarke, 2018; Samways et al., 2015). Four distinct phenotypic groups were identified, some of which corresponded to different sections of the Upper Salmon River (Clarke, 2018; Figure 1). Along with the genotypes of the juveniles, these morphological data serve as the foundation to assess the correlation between the phenotypic and genetic groupings.

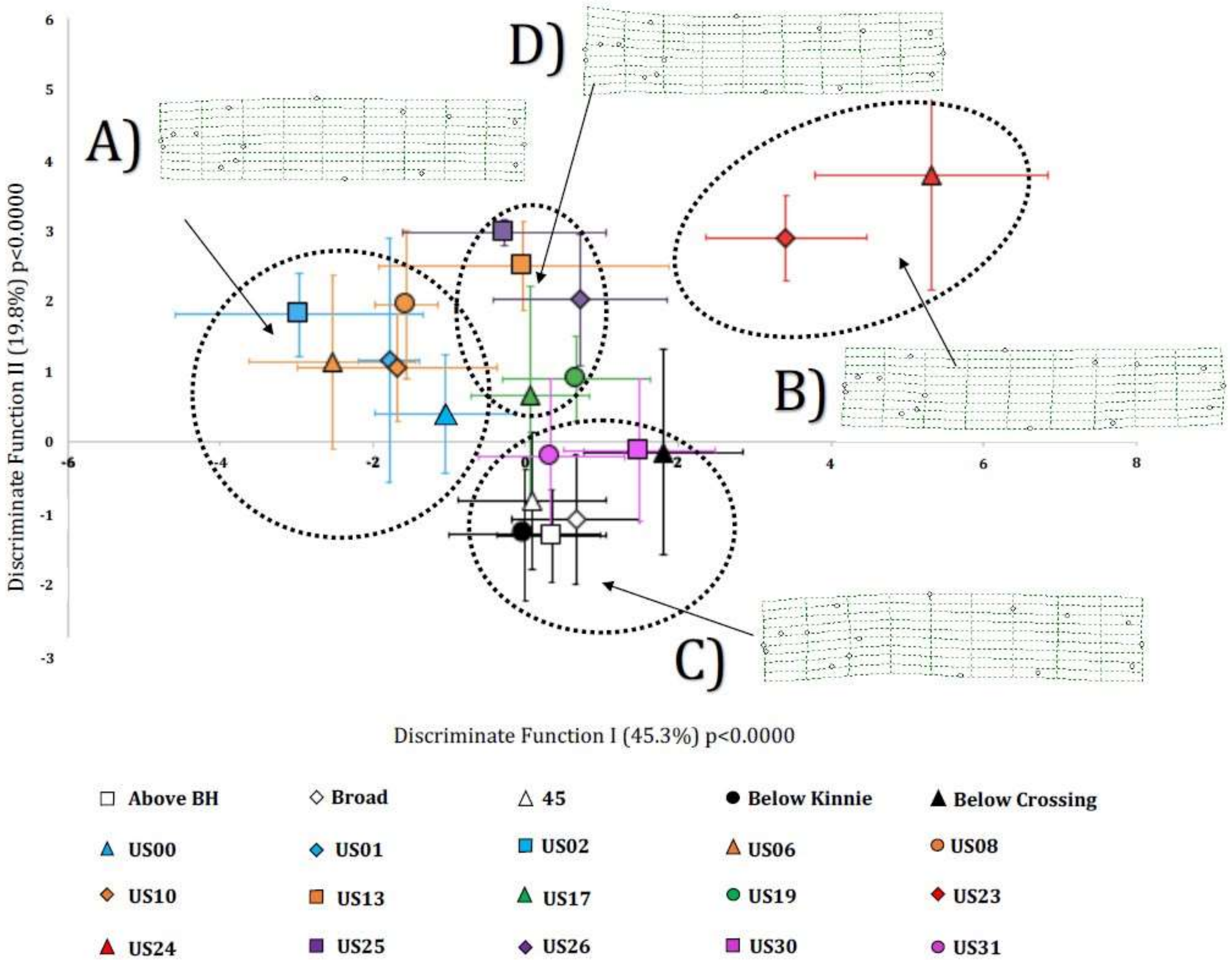

Figure 1: Discriminant function analysis (DFA) of Atlantic salmon young-of-the-year geometric morphometric data. Cluster membership for the four groups (A, B, C, D) is represented by the symbols and colours, surrounded by 95% confidence ellipses. The green matrix represents the relative shape differences between individuals within that group while the dots on the grid represent the biological landmarks used in the analysis. Figure is taken from Clarke, 2018.

Study site

The Upper Salmon River NB, Canada is situated in Fundy National Park with a drainage area of 174 km^2, and is a steep, deeply incised, fast flowing, mountainous tributary with deep pools and boulders that create shallow riffles and runs (Dadswell, 1968).

Most sampling sites are shallow, riffle habitats except for US23, US24, US25, US08, and some parts of US01 and US31, which are slightly deeper, with slower moving water, and are categorized as run habitat (Table 1). US23 is the deepest sampling site with a depth of approximately 42.3 cm, and US00 is the shallowest site with a depth of approximately 17.7 cm (Table 1). Mean water temperature in the ~6 km sampled stretch of the river was 18.8 ± 1.8 ℃ over the one-month sampling period (Table 1). There are no physical barriers within the river until Broad River and Match Factory Falls, situated on the Broad and 45 River branches, respectively (Figure 2).

Table 1: Site habitat data from sampling sites used for geometric morphometric and genetic analyses sampled in the summer of 2017 in the Upper Salmon River. Measured parameters include habitat type (riffle/run), water depth within the deepest part of the site (cm), and water temperature (°C).

Site	Habitat type	Water depth (cm)	Water temperature (°C)
US00	Riffle	17.67	19
US01	Riffle/Run	21.22	20
US02	Riffle	21.33	18
Broad	Riffle	N/A	N/A
45	Riffle	N/A	16
US06	Riffle	18.33	18
US08	Run	40.67	19
US10	Riffle	18.67	18
US13	N/A	39.33	21
US17	Riffle	24.67	18
US19	Riffle	22.67	20
Above BH	Riffle	N/A	22
US23	Run	42.33	22
US24	Run	20.67	19
US25	Run	38.67	19
US26	Riffle	22.33	17
Below Crossing	N/A	N/A	19
US30	Riffle	24.67	N/A
Below Kinnie	N/A	N/A	17
US31	Riffle/Run	19	16

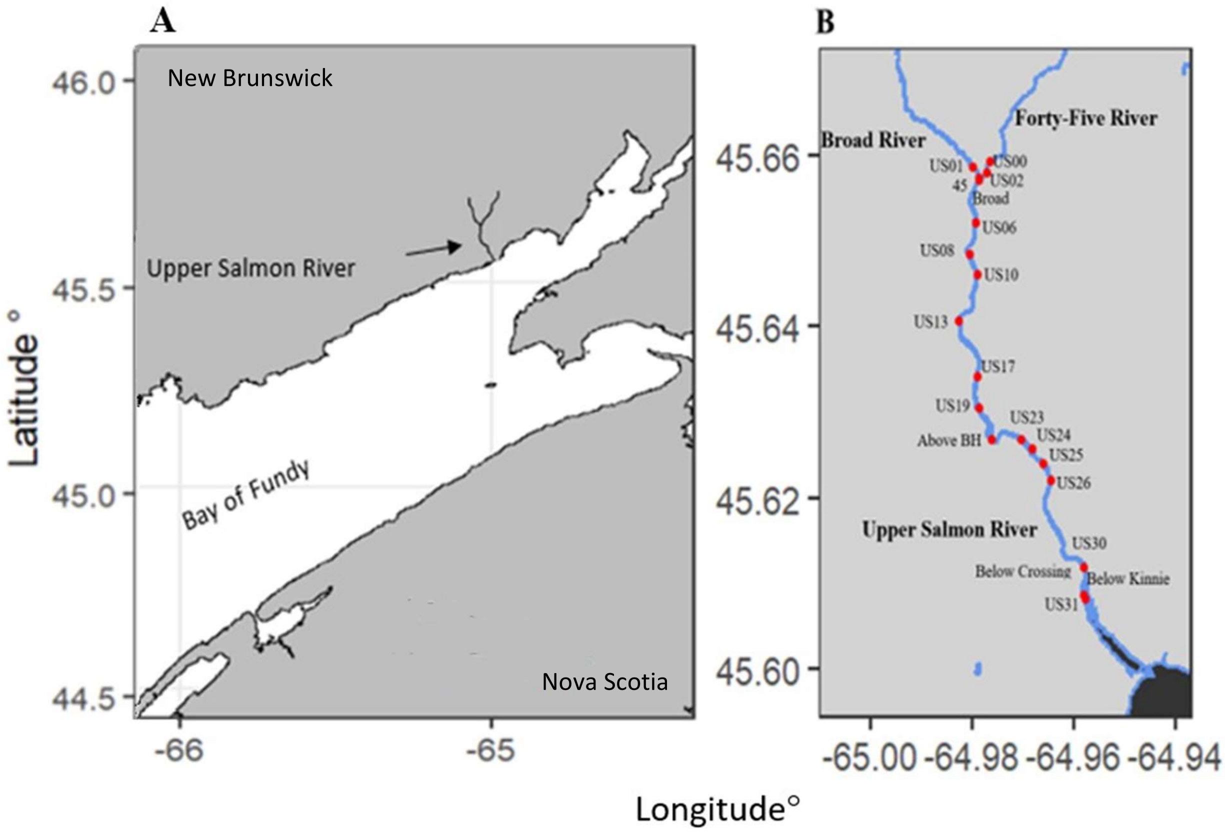

Figure 2: Map of the location of the Upper Salmon River (A) and the 20 sampling locations used for geometric morphometric and genetic analyses (B) represented by the red points.

Statistical analyses

Genetic cluster analyses

To determine whether YOY individuals within the Upper Salmon River fell within one or multiple genetic groupings, a set of genetic clustering analyses were conducted using discriminant analysis of principal components (DAPC). DAPC uses a multivariate, k-mean algorithm to maximize the between-population genetic differences while minimizing within-group differences (Jombart et al., 2010). DAPCs were conducted in R (V. 4.1.2) (R core team, 2021) using the package *adegenet* (Jombart, 2008).

A series of individual based DAPCs were conducted on the YOY salmon multi-locus genotypes. The first analysis used no prior population information (i.e., it did not include data on which location an individual was sampled from). Instead, the *find.clusters* function was used to determine the number of genetic clusters (k) within the YOY samples based off of the lowest Bayesian Information Criterion (BIC). After establishing there were six clusters, a DAPC was run with 100 principal components and five discriminant functions retained. To determine whether prior information about individuals' sampling locations and morphology yielded qualitatively similar results, additional DAPCs were conducted. The first used sampling site as a prior and considered up to 20 clusters (the number of sampling locations). The second analysis used the major regions of the river as a prior, testing up to three clusters. Here, sites within the river were combined as follows: Region 1: US00, US01, US02, Broad, 45, US06, US08, and US10, Region 2: US13, US17, US19, Above BH, US23, US24, US25, and US26, Region 3: Below Crossing, US30, Below Kinnie, and US31. The final analysis used the morphological grouping that the individuals fell within (A, B, C, D) from the geometric morphometric analysis by Clarke, (2018) as a prior, testing up to four clusters (Figure 5). This analysis tested whether individuals within a morphological grouping were also members of the same genetic cluster, based on this set of seven microsatellites.

Summary statistics

F-statistics are commonly used in population genetics to quantitatively estimate population structure using genetic markers (Goudet, 2005). Pairwise F_{ST} values were calculated as a measure of genetic differentiation between sampling sites. F_{ST} values between pairs of sampling sites were calculated in R using the '*pairwise.WCfst*' function from the package *heirfstat* (Goudet, 2005).

An analysis of molecular variance (AMOVA) was conducted using the *poppr* package in R (Kamvar et al., 2014) to partition genetic variance among different hierarchical levels. Specifically, genetic variance was portioned into the following levels: among-individual samples, between sampling sites within regions of the river, and between regions of the river.

Isolation by distance

Isolation by distance tests were run to determine whether the YOY individuals from a sampling location were more genetically dissimilar (less closely related) as the geographic distance between them increased. The geographic distance between pairwise sampling locations was calculated using Euclidean distance in R using the package *sp* (Pebesma and Bivand, 2005). The linearized pairwise F_{ST} matrix between sampling sites was regressed against this geographic distance matrix. To confirm the pattern from the regression, a correlative Mantel test was also run using 999 permutations and an alpha value of 0.05 using the package *vegan* (Oksanen et al., 2020).

Results

Hierarchical patterns of genetic structure

Clustering analyses revealed genetic structure within this ~6km stretch of the Upper Salmon River; however, the number of genetic groups was sensitive to whether location priors were used. The DAPC that used no prior assumptions indicated the presence of six genetic clusters (Figure 3). Although the lowest BIC suggested there were six total clusters, several groups were partially overlapping in 2-dimensional space (for example groups three and four or groups five and six; Figure 3). This indicates that when no prior information is given in the clustering analysis, four unique genetic subgroups is the best fit for the data, where individuals within a group are more genetically similar to each other than to individuals within another genetic group. The DAPC that used sampling site as a prior resulted in 20 genetic clusters, each of which corresponded to a unique sampling site (Figure 4). Many sites overlapped, however, sites that were closer together geographically were, on average, more genetically similar. For example, the sites that were lower in the river (closer to the mouth of the river) were more differentiated from the rest of the sites (Figure 4).

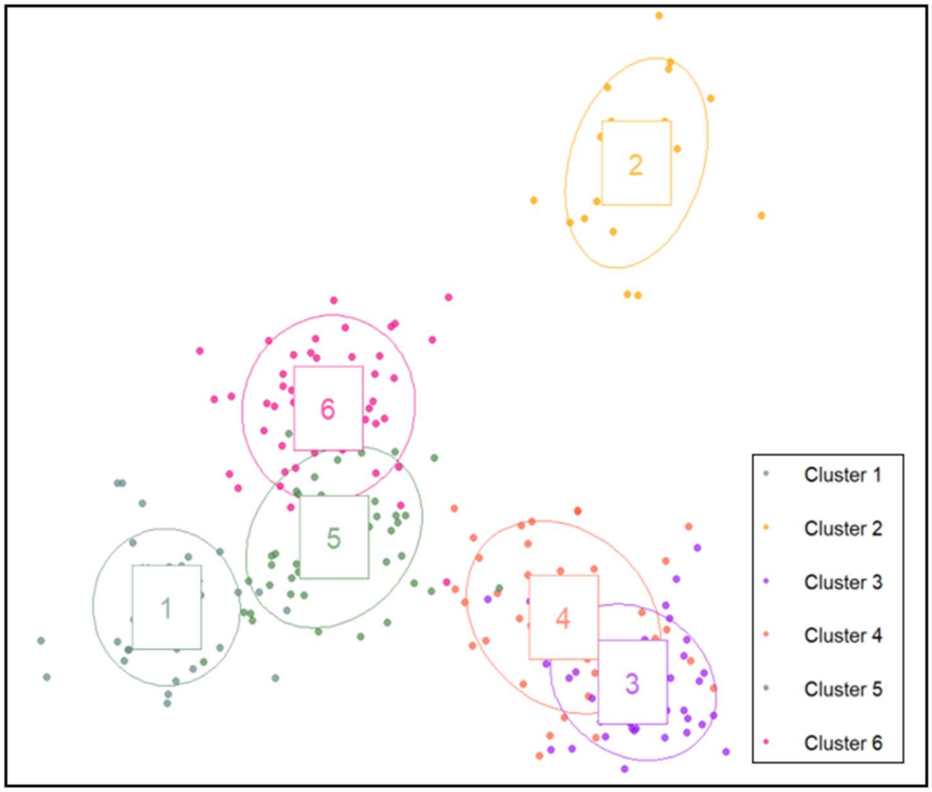

Figure 3: Discriminant analysis of principal components (DAPC) on young-of-the-year Atlantic salmon microsatellite genotypes (n=252 individuals) using no prior location data. Cluster membership is represented by the colours of points. Ovals represent 95% confidence ellipses.

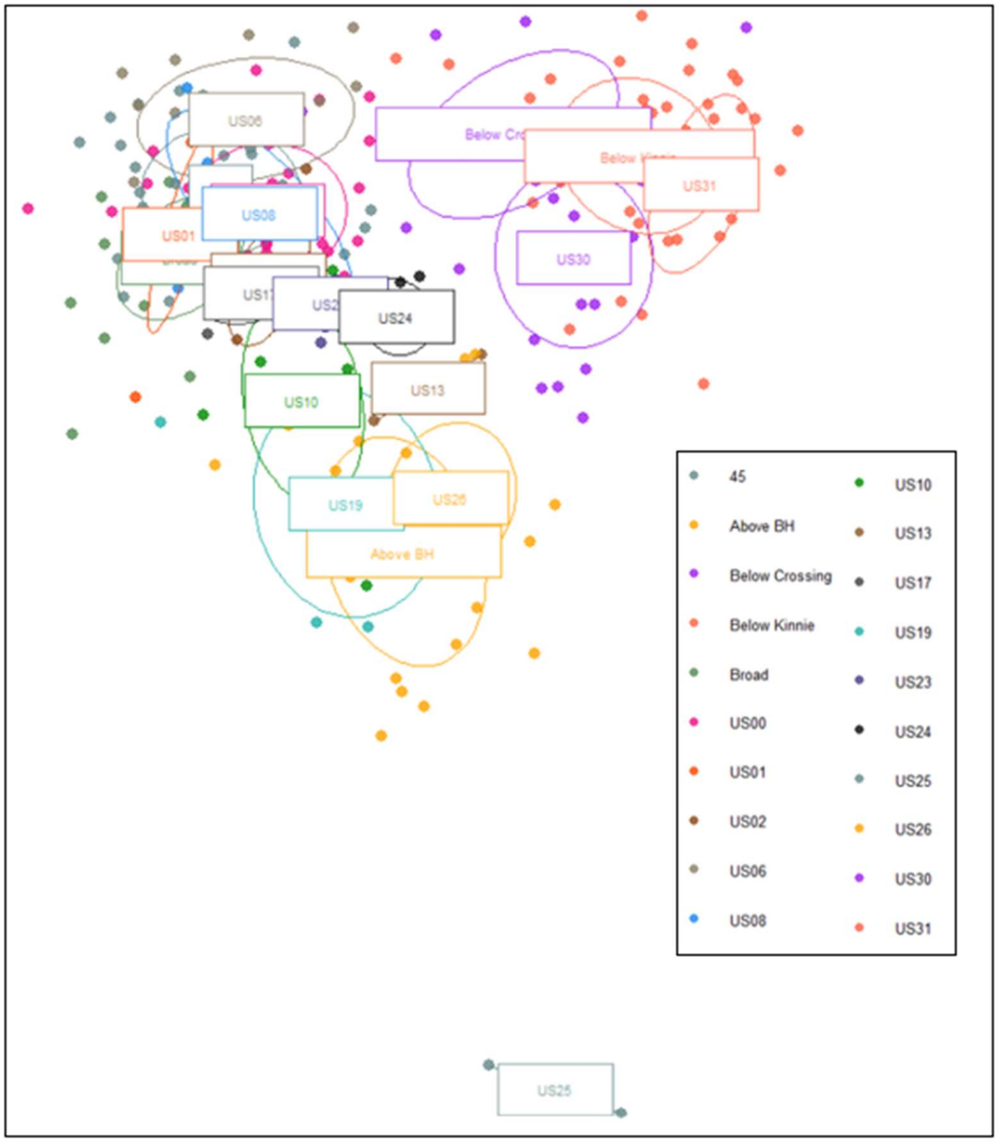

Figure 4: Discriminant analysis of principal components (DAPC) on young-of-the-year Atlantic salmon microsatellite genotypes (n=252 individuals) using sampling location (n=20) as prior knowledge. Cluster membership within sampling locations is represented by the colours of points. Ovals represent 95% confidence ellipses.

Because the initial DAPCs suggested regional structuring, sites were subsequently grouped by the region of the river that they fell within to visualize a clearer genetic delineation between geographically separated sites. The regions of the river were separated into three groupings, equivalent to the classification assigned by Parks Canada. These groupings included the upper most sites, farthest from the mouth of the river, the middle sites, and the lower sites closest to the mouth of the river, commonly referred to as the Forks, Black Hole, and Pumphouse regions respectively. The DAPC that used region of the river as a prior resulted in three distinct genetic clusters, each of which corresponded to a different region of the river (Figure 5). Overall, these results reveal signals of genetic clustering between regions of the river within only ~6 km.

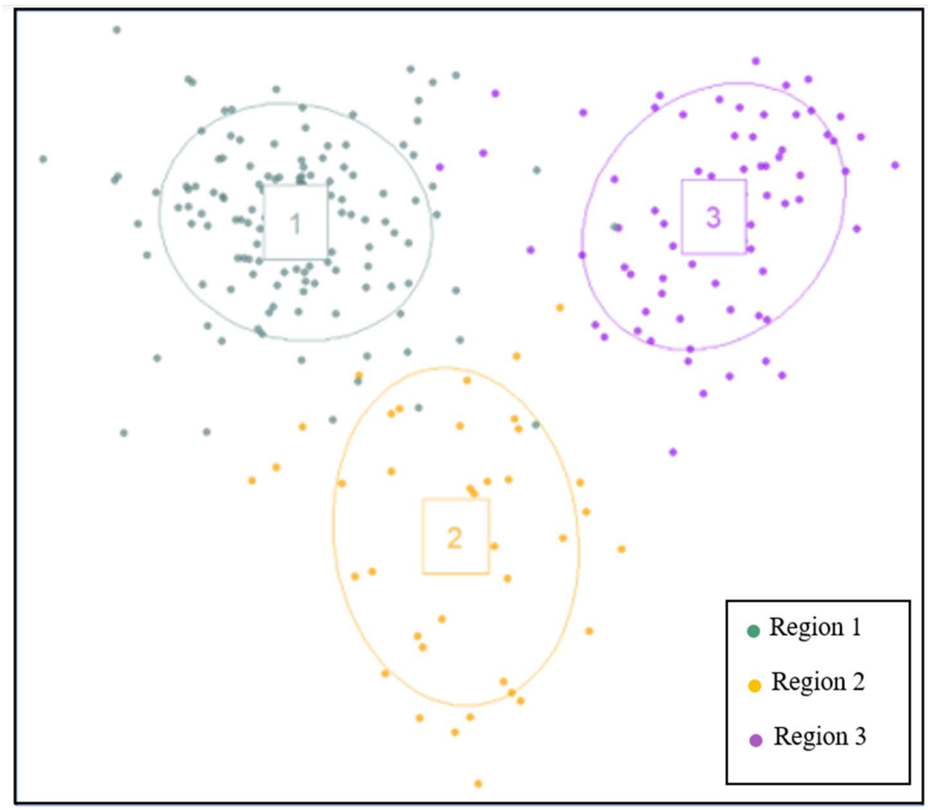

Figure 5: Discriminant analysis of principal components (DAPC) on young-of-the-year Atlantic salmon microsatellite genotypes (n=252 individuals) using geographic region of the river as prior knowledge. Cluster membership within regions of the river is represented by the colour of points. Ovals represent 95% confidence ellipses.

Analysis of molecular variance (AMOVA) results were congruent with the DAPC (Table 2). Specifically, they showed that, while most of the total genetic variance (92.43%) was explained at the individual level (i.e., by unique among-individual genotypes), a small amount of variance was explained by differences between sampling sites within region of the river (2.45%). Even after accounting for individual-individual and site-site variance, 5.12% of the variance was explained by differences between the three regions of the river (Table 2). Collectively, these results indicate that the extent of genetic differentiation between populations from different regional groups was approximately twice that observed between sampling sites within any region and indicates a stronger pattern of genetic structure at the regional level.

Table 2: Hierarchical analysis of molecular variance (AMOVA) of microsatellite loci allele frequencies among samples of young-of-the-year Atlantic salmon in the Upper Salmon River.

Source of variance	d.f.	Percentage of variance	*p*
Between regions of the river	2	5.12	< 0.001
Between sampling sites within region of the river	17	2.45	< 0.001
Within samples	232	92.43	< 0.001

Lastly, to test whether there was an association between the genetic and previously identified morphological groupings, a DAPC that used the morphological groupings as a prior was run. The DAPC revealed that some but not all morphological groupings were genetically distinct (Figure 6). Morphological group B and D appeared to be distinct genetic groups, but morphological groups A and C were largely overlapping (Figure 6). Overall, these results indicated that four morphological groupings corresponded best to roughly three genetic groupings.

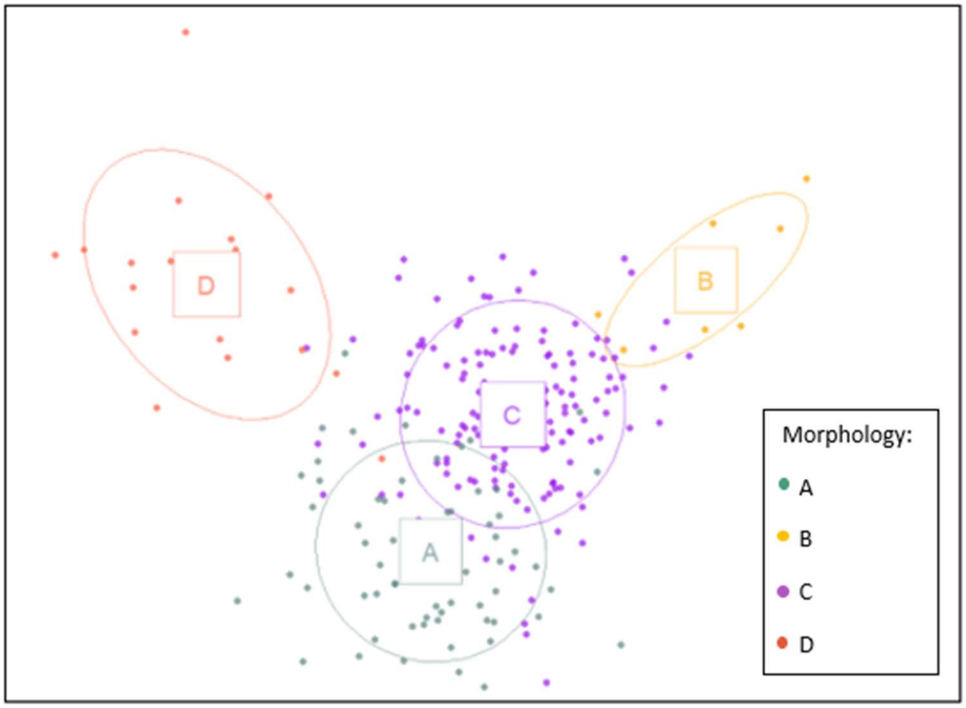

Figure 6: Discriminant analysis of principal components (DAPC) on young-of-the-year Atlantic salmon microsatellite genotypes (n=252 individuals) using morphological groupings as prior knowledge. Cluster membership is represented by the colour of points. Ovals represent 95% confidence ellipses.

Pairwise F_{ST} and isolation by distance

The mean F_{ST} value across all sites was 0.0360 ± 0.0286 which indicates a subtle level of genetic differentiation (Table 3). The pair of sites with the highest pairwise F_{ST} was between US13 and US24 (F_{ST} =0.1591). These sites are separated by a distance of 1.82 km and are located within the same region of the river (region 2). The lowest pairwise F_{ST} was 0 which occurred between multiple pairs of sites (n=17). The pairs of sites exhibiting no pairwise genetic structure (F_{ST} =0) were separated, on average, by a distance of 1.26 ± 1.01 km apart (Table 3).

Table 3: Matrix of pairwise F_{ST} estimates displaying the genetic difference in Atlantic salmon young-of-the-year between pairs of sampling sites of the Upper Salmon River.

	US00	US01	US02	Broad	45	US06	US08	US10	US13	US17	US19	Above BH	US23	US24	US25	US26	Below Crossing	US30	Below Kinnie
US00																			
US01	0.0227																		
US02	0.0263	0.0151																	
Broad	0.0197	0.0233	0.0131																
45	0.0065	0.0195	0.0298	0.0126															
US06	0.0000	0.0099	0.0062	0.0109	0.0000														
US08	0.0000	0.0000	0.0268	0.0100	0.0000	0.0000													
US10	0.0352	0.0495	0.0392	0.0465	0.0400	0.0167	0.0290												
US13	0.0442	0.0767	0.0528	0.0497	0.0512	0.0241	0.0270	0.0396											
US17	0.0288	0.0820	0.0803	0.0482	0.0208	0.0279	0.0092	0.0484	0.0152										
US19	0.0168	0.0227	0.0362	0.0199	0.0156	0.0137	0.0078	0.0190	0.0000	0.0174									
Above BH	0.0497	0.0418	0.0449	0.0529	0.0447	0.0282	0.0360	0.0704	0.0882	0.0997	0.0517								
US23	0.0780	0.0328	0.0987	0.1012	0.0650	0.0409	0.0359	0.0839	0.1591	0.1327	0.0921	0.0354							
US24	0.0134	0.0083	0.0178	0.0218	0.0124	0.0000	0.0000	0.0322	0.0571	0.0711	0.0483	0.0014	0.0000						
US25	0.0817	0.0232	0.0418	0.0851	0.0850	0.0450	0.0582	0.0000	0.0000	0.1072	0.0064	0.0287	0.0412	0.0469					
US26	0.0337	0.0070	0.0499	0.0539	0.0445	0.0184	0.0041	0.0239	0.0000	0.0590	0.0145	0.0564	0.0544	0.0251	0.0274				
Below Crossing	0.0362	0.0285	0.0602	0.0566	0.0431	0.0096	0.0235	0.0174	0.0570	0.0723	0.0274	0.0489	0.0657	0.0182	0.0538	0.0092			
US30	0.0318	0.0243	0.0356	0.0470	0.0367	0.0113	0.0152	0.0186	0.0176	0.0509	0.0165	0.0333	0.0393	0.0073	0.0288	0.0000	0.0000		
Below Kinnie	0.0461	0.0250	0.0592	0.0608	0.0519	0.0257	0.0325	0.0516	0.0634	0.0811	0.0420	0.0336	0.0507	0.0240	0.0564	0.0214	0.0047	0.0004	
US31	0.0567	0.0303	0.0643	0.0742	0.0641	0.0356	0.0383	0.0595	0.0591	0.1027	0.0529	0.0430	0.0479	0.0309	0.0407	0.0191	0.0144	0.0000	0.0000

Considering all pairs of sites, isolation by distance (IBD) analyses indicate that pairwise multi-locus F_{ST} estimates had a weak positive linear association with geographic distance between sampling sites (Figure 7; m= 0.0041, r^2=0.0436, $F_{1, 188}$=9.624, *P*= 0.0022). In other words, as the geographic distance between sites increased, so did the genetic distance between them. Although a statistically significant relationship between linearized F_{ST} and geographic distance was found, the fit of the linear model was poor with geographic distance, explaining a low amount of the overall variation in pairwise F_{ST}. This IBD pattern was also confirmed with a correlative Mantel test (Mantel r=0.2207, *P*=0.008).

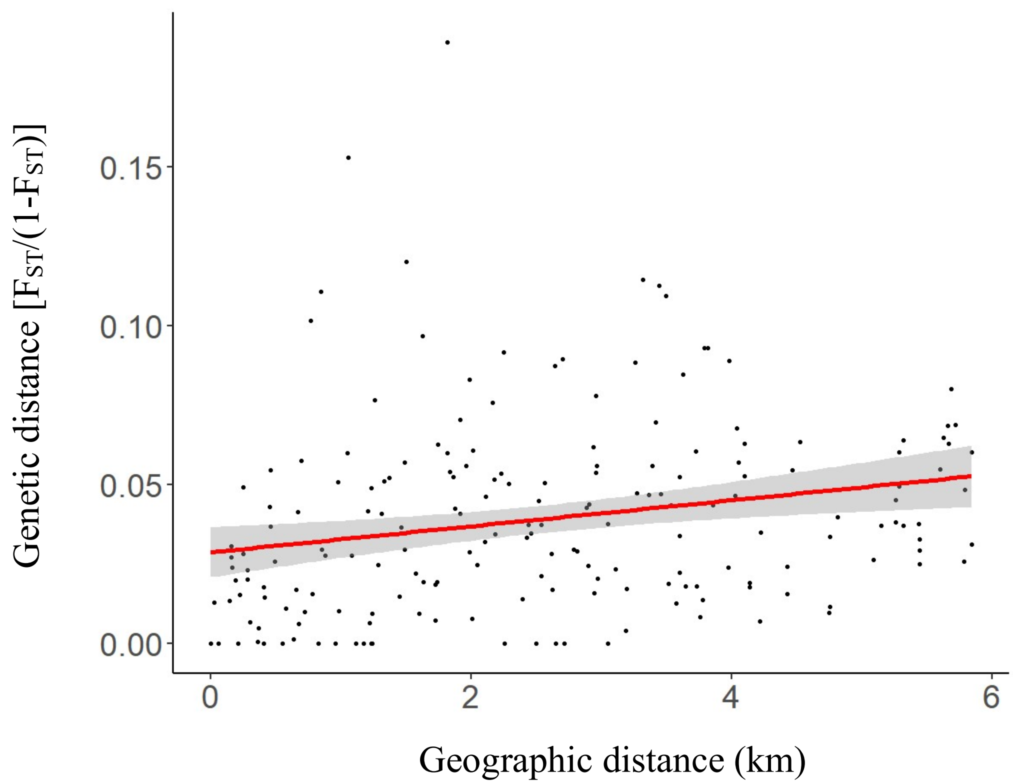

Figure 7: Isolation by distance linear regression depicting the association between pairwise linearized genetic distance [F_{ST}/(1-F_{ST})] and increasing geographic isolation (km) between sites where young-of-the-year Atlantic salmon were sampled. Grey shading represents 95% confidence intervals.

Discussion

The purpose of this study was to determine the population genetic structure of YOY Atlantic salmon within the Upper Salmon River as well as to explore the genetic similarity among previously defined morphological groups using clustering analyses of individuals' microsatellite genotypes. The present findings suggest that there is fine scale population structure of YOY Atlantic salmon within the Upper Salmon River. This is an interesting finding considering the small spatial scale of the study (~6 km), and the drastic decrease in population size and genetic diversity of inner Bay of Fundy Atlantic salmon over the past few decades. Results also indicate that there may be some genetic similarity between previously defined morphological groups, suggesting at least a partial genetic component to the observed phenotypic variation. However, the lack of a strong genetic correlation to all phenotypic groups indicates a potential environmental component to phenotypic variation as well. The overall findings suggest that the processes driving genetic structure may be acting at a smaller scale than typically thought, within small watersheds or tributaries, and conservation design at this scale may be important to consider as part of management strategies.

Potential drivers of fine-scale genetic structure

Hierarchical genetic structure exists within this ~6 km stretch of the Upper Salmon River. Pairwise F_{ST} estimates increased linearly with increasing geographic distance but was a poor fit to the data, indicating a weak positive isolation by distance (IBD) pattern along the river. This type of IBD pattern reflects spatially restricted gene flow (Wright, 1942; Malécot 1948). Although IBD is generally a common pattern of spatial genetic structure in marine fishes (including anadromous fishes) (Selkoe et al., 2016), IBD at this geographic scale is surprising and has rarely been documented in salmonid populations. A few studies have shown an IBD pattern at a similar spatial scale in salmonids that also display homing behaviour such as masu salmon over 21 km (Kitanishi et al., 2009), or brown trout over less than 10 km (Carlsson et al., 1999). Atlantic salmon populations have been documented to display IBD over much greater spatial scales of 100's of kilometers in distance between sites or between watersheds (King et al., 2001; Dillane, 2007; Dionne et al., 2008; Tonteri et al., 2009) except for Primmer et al., (2006) in which weak but significant IBD was observed over a geographic scale of approximately 11 km.

It has been suggested that natal homing is a primary driver of fine-scale population structure in many anadromous fish species since individuals have been observed to home to the tributary level (Stabell, 1984; Vähä et al., 2007). Because of the high fidelity of individuals returning to their spawning grounds, homing may act as a barrier to gene flow even within a watershed with no physical barriers to dispersal. This is likely the primary mechanism explaining population structure within the Upper Salmon River in the present study since there are no physical barriers along the sampled river area; however, there seems to be a weak, yet significant linear IBD pattern. While it is possible that these patterns are also influenced by local adaptation, there is little environmental variability over the sampled area. Moreover, we do not have the genetic data required to test for selection (see "Limitations of a small microsatellite marker panel" below).

Genetic association to identified phenotypic groupings

The second objective of this study, to determine the genetic correlation to previously identified morphologies, remains inconclusive with the available findings. DAPC analyses revealed there was not a clear relationship between morphological groupings and genetic groupings, i.e., the individuals categorized in the four morphological groups (Clarke, 2018) did not correspond to distinct genetic clusters. Instead, only morphological groups B and D appeared to represent distinct genetic groupings. These results indicate that variation in phenotype may be influenced by either genetic differences which code for certain morphologies and can be passed on to offspring, or genetic differences that arise due to local adaptation from differing selective pressures within microenvironments along the river. Morphological group A and C are distinct morphologies; however, they were found to be very similar genetically (belong to one genetic grouping). These results indicate that genetics alone could not have been the primary driver of the variation in phenotype, and there may have been more of an environmental influence involved. A different pattern may be observed when testing different regions of the genome, however, and sets forth the opportunity for future research. These findings partially support my prediction that fish within a morphological grouping would be more genetically similar to each other than to fish in another morphological grouping since individuals with morphology B and D were more genetically similar within groups (i.e., distinct genetic groupings). However, this prediction was not fully supported since morphologies A and C were as genetically similar between groups as they were within groups (high degree of overlap between genetic groups).

The driver(s) of the morphologies cannot be directly tested with the available data. To test for heritability of key morphological traits (e.g., length, weight, condition factor, centroid size), genome-wide single nucleotide polymorphisms (SNP) data would be needed (Speed et al., 2012) as well as information on familial relationships, obtained through a pedigree or genetic relatedness analyses. Additionally, with more extensive environmental data on the sampled river area, more robust conclusions can be drawn in terms of the potential influence of phenotypic plasticity.

Limitations of a small microsatellite marker panel

Identifying the spatial scale at which genetic structuring occurs is a major goal of population genetics since it is key to identifying genetic units for conservation (Primmer et al., 2006). The use of microsatellite markers is a relatively inexpensive method for determining genetic relationships at the population and individual level (Primmer et al., 2006). Microsatellite loci have a high mutation rate which increases the rate of population divergence. Because of this, they are often preferred over other traditional markers, such as mitochondrial DNA where there may be limited variability and limited detection of divergence (Angers and Bernatchez, 1998). However, due to the neutrality of these markers, adaptive effects are not able to be quantified, limiting the conclusions that can be drawn from the data. Additionally, when there are a low number of markers in the genetic panel being used, genetic relatedness values are inflated (D'Aloia et al., 2018). This issue is particularly problematic for parentage and sibship analyses, but it is possible that a small marker panel could bias the genetic clustering analyses used here. This could have influenced the results of my study since the data was a microsatellite marker panel with only seven loci, thus these results must be interpreted with caution.

Conservation implications

Population genetic processes play a role in conservation at many levels. Determining the spatial scale at which genetic structuring occurs is the first step to understanding ecological and evolutionary drivers of genetic differentiation patterns (Kitanishi et al., 2009; Dionne et al., 2009; Tabatabaei et al., 2020). It is also important to quantify the patterns of connectivity for severely depressed and isolated populations to determine if a population can serve as a source for recovery within a metapopulation (Kelson et al., 2015).

Many Atlantic salmon conservation initiatives have focused on captive reared salmon to mitigate marine mortality through the re-introduction of wild origin salmon fry to their natal rivers (King et al., 2001; O'Sullivan et al., 2020). The iBoF Atlantic salmon population has less than 225 wild origin individuals left and has avoided extinction through these captive rearing programs (e.g., DFO Live Gene Bank; Fundy Salmon Recovery) since the 1990's (DFO, 2010). Out of the ~40 historic iBoF Atlantic salmon rivers, only six have such conservation initiatives actively releasing salmon of various life stages (DFO, 2010). This low proportion of conserved historic rivers (~15%) may not be sufficient for the recovery of the population. In hopes of recovering the population, conservation initiatives will need to include more rivers since natural straying from natal rivers is relatively low (<12%) and will likely not contribute enough to gene flow to re-populate all historic iBoF salmon rivers (DFO, 2010).

One important consideration for these conservation initiatives is quantifying whether there is a shift in the gene frequencies within natal rivers, and the amount of gene flow occurring between geographic populations (King et al., 2001). Within-river genetic processes are also important to consider with respect to salmon aquaculture since captive salmon escapees have been seen to disturb the evolutionary processes of some wild populations due to interbreeding of wild and captive salmon or through the introduction of disease, leading to decreased fitness of wild populations (Jensen et al., 2010). Additionally, determining the stage at which local adaptation is occurring within populations can be used to indicate how anthropogenic management actions effect the behaviour and fitness of the individuals over time (King et al., 2001).

Establishing the proper spatial scale at which management action should be taken is important to avoid over- or underestimating extinction risks (Dionne et al., 2009). For example, if a river-based management approach is used when high within-river genetic structure exists, this could lead to an underestimation of extinction risks by assuming a larger census population size within the tributary (Dionne et al., 2009). However, over-estimating the within-river spatial genetic structure would lead to unnecessary management initiatives focused only on genetic structure rather than conservation of the population as a whole, as well as an under-estimation of connectivity between parts of the river or tributaries (Dionne et al., 2009). This is likely not an issue in the Upper Salmon River due to its small spatial scale and the lack of strong hierarchical genetic structure for these concepts to have a prominent influence on the extinction risk of the population. Additionally, because the iBoF population is protected under SARA, no fishing practices (commercial, recreational, or subsistence) are allowed, and many individuals are tagged to allow for a very accurate calculation of population size within each river (Marshall, 2013). However, these factors are important to consider for larger rivers where census population size may be hard to estimate and for species that are at risk from over-harvesting or environmental stressors. To reduce the chance of incorrectly interpreting the spatial scale of genetic structure, it has been suggested that restoration efforts take into consideration both inter- and intra-population genetic diversity (Fraser, 2008).

The Upper Salmon River is one of the focal rivers of the Fundy Salmon Recovery program (FSR), an innovative restoration initiative aimed at restoring the iBoF Atlantic salmon population (Bryson et al., 2022). The FSR program is based on a novel conservation strategy to collect out-migrating smolts, rear them in the marine environment in modified aquaculture sea cages at the world's first Marine Conservation Farm on the Island of Grand Manan in New Brunswick, and subsequently reintroduce mature adult salmon back to their natal river (Bryson et al., 2022). The FSR model differs from traditional captive brood stock programs in that sexual and natural selection determines offspring production instead of artificial and domestic influences (Bryson et al., 2022). The results of this study will serve as a baseline for potential genetic and ecological processes occurring within the Upper Salmon River that can be used by the FSR program to assess the most representative number of conservation units within the river. This information can be implemented by FSR to actively manage their program to mitigate any potential loss of genetic diversity.

Although only subtle genetic structure was observed, the results indicate that there is higher genetic variance explained between regions of the river than between sites within the regions of the river. This may imply that there are three potential management units within the river that correspond to the upper, middle, and lower regions of the river respectively. It has previously been acknowledged by Potter et al., (2003) and Youngson et al., (2003) that river-based management may be an oversimplification of Atlantic salmon life history and that a more precautionary management approach may be better suited for river systems. This could apply to the results found in this study if the population is managed at the regional level, but this approach may be more applicable to rivers/tributaries of a larger geographic scale with more apparent hierarchical genetic structure between subgroups.

Conclusion

In summary, this study demonstrates that population structure exists within the Upper Salmon River at a surprisingly small scale (~6 km) and that at least three separate management units may be considered. There are also potentially genetic and environmental influences of morphology that requires further research to determine the exact mechanism. The results from this study provide an additional line of evidence to consider adapting river-based management strategies to maintain the genetic structure of the population within the Upper Salmon River. Overall, this study highlights the importance of observing the genetic substructure of smaller spatial scales, within tributary systems, to capture both intra- and inter-river genetic diversity to make more informed management decisions.

References

Álvarez-Varas, R., Heidemeyer, M., Riginos, C., Benítez, H.A., Reséndiz, E., Lara-Uc, M., Godoy, D.A., Muñoz-Pérez, J.P., Alarcón-Ruales, D.E., Vélez-Rubio, G.M., Fallabrino, A., Piovano, S., Alfaro-Shigueto, J., Ortiz-Alvarez, C., Mangel, J.C., Esquerré, D., Zárate, P., Medrano, C., León Miranda, F., Guerrero, F., Vianna, J.A., Véliz, D., 2021. Integrating morphological and genetic data at different spatial scales in a cosmopolitan marine turtle species: Challenges for management and conservation. Zoological Journal of the Linnean Society 191, 434–453.

Angers, B., Bernatchez, L., 1998. Combined use of SMM and non-SMM methods to infer fine structure and evolutionary history of closely related brook charr (*Salvelinus fontinalis*, Salmonidae) populations from microsatellites. Mol. Biol. Evol. 15, 143–159.

Asche, F., Bjerndal, T., Salvanes, K.G., 2005. The demand for salmon in the European Union: the importance of product form and origin. Agricultural Economics 46(1), 69-81.

Beland, K. F., A. T. Bielak., 2002. Atlantic salmon fisheries in eastern North America: the prince and the pauper. American Fisheries Society, 61-76.

Bradshaw, A.D., 1965. Evolutionary significance of phenotypic plasticity in plants. Advances in Genetics 13, 115–155.

Bryson, G., Kidd, K., Samways, K., 2022. Food web incorporation of marine-derived nutrients after the reintroduction of endangered inner Bay of Fundy Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences. (in *press*).

Bull, J.J., 1987. Evolution of phenotypic variance. Evolution 41(2), 303-315.

Carlsson, J., Olsen, K., Nilsson, J., Øverli, Ø., Stabell, O., 1999. Microsatellites reveal fine-scale genetic structure in stream-living brown trout. Journal of Fish Biology 55, 1290-1303.

Clarke, L., 2018. Using geometric morphometrics assess phenotypic variation in wild-hatched juvenile Atlantic Salmon (*Salmo salar*). University of New Brunswick. Unpublished document.

Dadswell, M.J., 1968. Atlantic salmon, *Salmo salar*, investigations in the Pointe Wolfe and Upper Salmon Rivers, Fundy National Park. Canadian Wildlife Service: Limnology Section. 1-53.

D'Aloia, C.C., Xuereb, A., Fortin, M.J., Bogdanowicz, S.M., Buston, P.M., 2018. Limited dispersal explains the spatial distribution of siblings in a reef fish population. Marine Ecology Progress Series 607, 143–154.

Davidson, W., Birt, T., Green, J., 1989. A review of genetic variation in Atlantic salmon. Journal of Fish Biology 34, 547–560.

DFO. 2010. Recovery strategy for the Atlantic salmon (*Salmo salar*), inner Bay of Fundy populations [online]. [Final.] Species at Risk Act Recovery Strategy Series, Fisheries and Oceans Canada, Ottawa, Ontario.

Dillane, E., 2007. Spatial and temporal patterns in microsatellite DNA variation of wild Atlantic salmon, *Salmo salar*, in Irish rivers. Fisheries Management and Ecology 14, 209-219.

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

Dionne, M., Caron, F., Dodson, J.J., Bernatchez, L., 2009. Comparative survey of within-river genetic structure in Atlantic salmon; Relevance for management and conservation. Conservation Genetics 10, 869–879.

Fraser, D.J., 2008. How well can captive breeding programs conserve biodiversity? A review of salmonids. Evolutionary Applications 1, 535–586.

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

Fusco, G., Minelli, A., 2010. Phenotypic plasticity in development and evolution: Facts and concepts. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 547–556.

Garant, D., Dodson, J., Bernatchez, L., 2000. Ecological determinants and temporal stability of the within-river population structure in Atlantic salmon (*Salmo salar* L.). Molecular Ecology 9, 615-628.

Gentili, R., Solari, A., Diekmann, M., Duprè, C., Monti, G.S., Armiraglio, S., Assini, S., Citterio, S., 2018. Genetic differentiation, local adaptation and phenotypic plasticity in fragmented populations of a rare forest herb. PeerJ, 1-26.

Goudet, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. Molecular Ecology Notes 5, 184–186.

Hallgrimsson, B., Hall, B.K., 2005. Variation: A Central Concept in Biology. Elsevier Academic Press.

Hansen, L., Jonsson, N., Jonsson, B., 1993. Oceanic migration in homing Atlantic salmon. Animal Behaviour 45, 927–941.

Haugen, T.O., Vollestad, L.A., 2000. Population differences in early life-history traits in grayling. Journal of Evolutionary Biology 13, 897-905.

Jensen, A.J., Hansen, L.P., Johnsen, B.O., Karlsson, S., 2017. Rapid evolution of genetic and phenotypic divergence in Atlantic salmon following the colonisation of two new branches of a watercourse. Genetics Selection Evolution 49, 1–12.

Jensen, O., Dempster, T., Thorstad, E., Uglem, I., Fredheim, A., 2010. Escapes of fishes from Norwegian sea-cage aquaculture: Causes, consequences and prevention. Aquaculture Environment Interactions 1, 71-83.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403-1405.

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. BMC Genetics 11.

Kamvar, Z., Tabima, J., Grünwald, N., 2014. Poppr: a R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2.

Kelson, S.J., Kapuscinski, A.R., Timmins, D., Ardren, W.R., 2015. Fine-scale genetic structure of brook trout in a dendritic stream network. Conservation Genetics 16, 31–42.

King, T.L., Kalinowski, S.T., Schill, W.B., Spidle, A.P., Lubinski, B.A., 2001. Population structure of Atlantic salmon (*Salmo salar* l.): A range-wide perspective from microsatellite DNA variation. Molecular Ecology 10, 807–821.

Kitanishi, S., Yamamoto, T., Higashi, S., 2009. Microsatellite variation reveals fine-scale genetic structure of masu salmon, *Oncorhynchus masou*, within the Atsuta River. Ecology of Freshwater Fish 18, 65–71.

Lucek, K., Sivasundar, A., Seehausen, O., 2014. Disentangling the role of phenotypic plasticity and genetic divergence in contemporary ecotype formation during a biological invasion. Evolution 68, 2619–2632.

Malécot, G., 1948. Mathématiques de l'hérédité. Paris: Masson et Cie.

Marshall, T.L., 2013. Inner Bay of Fundy (iBoF) Atlantic Salmon (*Salmo salar*) marine habitat: Proposal for important habitat. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/071. vi + 69 p.

Oostra, V., Saastamoinen, M., Zwaan, B.J., Wheat, C.W., 2018. Strong phenotypic plasticity limits potential for evolutionary responses to climate change. Nature Communications 9, 1-11.

Oksanen, J., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P., O'Hara, R., Simpson, G., Solymos, P., Stevens, M., Szoecs, E., Wagner, H., 2020. vegan: Community ecology package. R package version 2.5-7.

O'Sullivan, R.J., Aykanat, T., Johnston, S.E., Rogan, G., Poole, R., Prodöhl, P.A., de Eyto, E., Primmer, C.R., McGinnity, P., Reed, T.E., 2020. Captive-bred Atlantic salmon released into the wild have fewer offspring than wild-bred fish and decrease population productivity. Proceedings of the Royal Society B: Biological Sciences 287, 16-71.

Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics 25, 547–572.

Potter, E. C. E., Wyatt, R., McLean, J., and Campbell, R., 2003. Managing the exploitation of migratory salmonids. Fisheries Research 62, 127-142.

Pebesma, E.J., Bivand. R.S., 2005. Classes and methods for spatial data in R. R News 5(2).

Prevost, A., Fraser, D., 2018. DNA Parentage analysis of Atlantic salmon adult rearing groups in Fundy National Park. Parks Canada, 1–16.

Price, T.D., Qvarnström, A., Irwin, D.E., 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society B: Biological Sciences 270, 1433–1440.

Primmer, C.R., Veselov, A.J., Zubchenko, A., Poututkin, A., Bakhmet, I., Koskinen, M.T., 2006. Isolation by distance within a river system: Genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. Molecular Ecology 15, 653–666.

R Core Team., 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Samways, K.M., Leavitt, P.R., Magnan, P., Rodríguez, M.A., Peres-Neto, P.R., 2015. Convergent polymorphism between stream and lake habitats: The case of brook char. Canadian Journal of Fisheries and Aquatic Sciences 72, 1406–1414.

Schmid, M., Guillaume, F., 2017. The role of phenotypic plasticity on population differentiation. Heredity 119, 214–225.

Selkoe, K.A., D'Aloia, C.C., Crandall, E.D., Iacchei, M., Liggins, L., Puritz, J.B., von der Heyden, S., Toonen, R.J., 2016. A decade of seascape genetics: Contributions to basic and applied marine connectivity. Marine Ecology Progress Series 554, 1–19.

Speed, D., Hemani, G., Johnson, M.R., Balding, D.J., 2012. Improved heritability estimation from genome-wide SNPs. American Journal of Human Genetics 91, 1011–1021.
Stabell, O., 1984. Homing and olfaction in salmonids: A critical review with special reference to the Atlantic salmon. Biol. Rev. 59, 333–388.
Tabatabaei, S.N., Abdoli, A., Hashemzadeh Segherloo, I., Normandeau, E., Ahmadzadeh, F., Nejat, F., Bernatchez, L., 2020. Fine-scale population genetic structure of Endangered Caspian Sea trout, *Salmo caspius*: implications for conservation. Hydrobiologia 847, 3339–3353.
Taylor, E.B., 1991. A review of local adaptation in Salmonidae, with particular reference to Pacific and Atlantic salmon. Aquaculture 98, 185–207.
Tonteri, A., Veselov, A.J., Zubchenko, A. v., Lumme, J., Primmer, C.R., 2009. Microsatellites reveal clear genetic boundaries among Atlantic salmon (*Salmo salar*) populations from the Barents and White seas, northwest Russia. Canadian Journal of Fisheries and Aquatic Sciences 66, 717–735.
Trzcinski, K.M., Gibson, J.F., Amiro, P.G., Randall, R.G., 2004. Inner Bay of Fundy Atlantic salmon (*Salmo salar*) critical habitat case study. Canadian Science Advisory Secretariat (CSAS) Fisheries and Oceans Canada.
Vähä, J.P., Erkinaro, J., Niemelä, E., Primmer, C.R., 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. Molecular Ecology 16, 2638–2654.
Vincent, B., Dionne, M., Kent, M.P., Lien, S., Bernatchez, L., 2013. Landscape genomics in Atlantic salmon (*Salmo salar*): Searching for gene-environment interactions driving local adaptation. Evolution 67, 3469–3487.
Wang, J., 2018. User's guide for software COLONY Version 2.0.6.5 5, 1–72.
Willem, J., Ferguson, H., 2002. On the use of genetic divergence for identifying species. Biological Journal of the Linnean Society 75, 509–516.
Wright, S., 1943. Isolation by distance. Genetics 28(2), 114- 138.
Youngson, A.F., Jordan, W.C., Verspoor, E., Cross, T., McGinnity, P., Ferguson, A., 2003. Management of salmonid fisheries in the British Isles: towards a practical approach based on population genetics. Fisheries Research 62, 193-209.