ECOLOGICAL SPECIATION OF A LANDLOCKED POPULATION OF RAINBOW SMELT (*Osmerus mordax*) IN LOCHABER LAKE, NOVA SCOTIA

by

Julie Rivard

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
June 2016

DEDICATION PAGE

For my parents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The process of speciation still remains one of the most debated issues in evolutionary biology. Studying sympatrically speciated pairs of fish, found in post-glacial lakes, may provide a unique insight into speciation, as populations are recently diverged. A landlocked sympatric pair of Rainbow Smelt (*Osmerus mordax*) found in Lochaber Lake, Nova Scotia is diverged into a small microphagous form and a large macrophagous form. Both forms have phenotypic differences associated with different trophic niches. Besides being phenotypically diverged they are also genetically differentiated suggesting reproductive isolation that is stable over time; however, results from this study provide evidence for gene flow between the two morphs. Hybrids were also observed in the adult and the larval Rainbow smelt providing further evidence that gene flow between the two morphs does exist, despite being selected against.

LIST OF ABBREVIATIONS USED

Abbreviations	Definition
A_E	Allelic richness
ANOVA	Analysis of Variance
BC	British Columbia
BSC	Biological Species Concept
CI	Confidence Interval
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide-triphosphate
F_{IS}	Inbreeding coefficient
FL	Fork length
F_{ST}	Genetic differentiation
He	Expected heterozygosity
Ho	Observed heterozygosity
HWE	Hardy Weinberg Equilibrium
LD	Linkage disequilibrium
MCMC	Markov Chain Monte Carlo
mtDNA	Mitochondrial DNA
N	Number of samples
N_A	Number of alleles
N_e	Effective population size
NB	New Brunswick
NS	Nova Scotia
PCoA	Principle coordinates analysis
PCR	Polymerase chain reaction
PSC	Phylogenetic Species Concept
SNP	Single nucleotide polymorphism
USA	United States of America

ACKNOWLEDGEMENTS

This thesis was completed under the supervision of Dr. Paul Bentzen. I thank him for his advice and guidance over the last several years. I thank fellow member of the Marine Gene Probe Lab, Ian Paterson for sharing his knowledge of population genetics and answers to general lab questions which allowed for proper data collection.

For help with sampling I thank Paul Bentzen, Ian Bradbury, Gregory McCracken, Meghan McBride, Vicky Yaroshewski, Devon Johnstone and Christopher Glibbery. Without them helping out with sample collection during the cold dark nights this project would not have been possible. I thank Ian Bradbury and Lynn Anstey for their previous work that provided already genotyped data.

Finally, I would like to thank my family, my friends and my husband, for their endless support and encouragement.

CHAPTER 1 INTRODUCTION

1.1 The study of adaptation and ecological speciation

The processes by which population divergence, and eventually speciation, occur remain among the central and most debated subjects in evolutionary biology (Taylor, 1999; Schluter, 2000; Barton, 2001; Saint-Laurent *et al*., 2003). This attention arises from the need to have taxonomic categorization of all living things, and because 'species', and not higher-level classifications, represent the taxonomic unit with the most substantial association to individuals and populations, which are central components of biodiversity (Taylor, 1999). The nature of species is very important in that the definition of what makes a species gives taxonomists the power to infer higher units of evolutionary diversity (Levin, 1979; Taylor 1999). Also, a framework for the features of species establishes relationships within and among clusters of individuals while characterizing the limits of what makes up a species. These associations and their limits therefore influence how we think about and study them (Taylor, 1999).

Despite the importance of species and speciation, concrete agreement on what specifically defines a species is still lacking (Barton, 2001). There are three criteria that are considered when trying to define species. Ideally, a species concept would be general, but applicable and theoretically significant (Hull, 1997). However, many of the criteria that a species concept should meet tend to conflict with one another, meaning one goal or criterion can only be realized at the expense of another (Hull, 1997). If a concept is easily applicable, it also tends to be theoretically trivial, but if a concept is theoretically significant, it will likely be difficult to apply (Hull, 1997). The debate on what defines a species has generated a number of concepts, all of them hoping to capture theoretical and/or operational aspects of a well thought out and relevant concept (Mayden, 1997). To date, there are approximately 22 species concepts

that have been proposed as definitions for speciation (Mayden, 1997; Taylor, 1999).

Of the 22 definitions however, only two are widely supported as operational concepts (Mayden, 1997; Taylor, 1999): the phylogenetic species concept (PSC; Cracraft, 1983) and the biological species concept (BSC; Mayr, 1940). The PSC encompasses at least three versions, but these all agree that species are the 'tip' on a phylogeny. In other words, the smallest group of organisms that share a common ancestor and can be distinguished from other such groups is considered a species (Mayden, 1997; Taylor, 1999). As for the BSC, which was described by Mayr (1940), species are defined as groups of individuals that can actually, or potentially, interbreed in nature and produce viable offspring but are reproductively isolated from other such groups.

The definition of the BSC has since been modified by Taylor (1999) to include certain cases where the potential for hybridization occurs between individuals of different species; however, the populations to which these individuals belong remain as distinct gene pools. Barton (2001) also added to the definition of the BSC whereby 'speciation' consists of the evolution of reproductive isolating mechanisms that result in the prevention of interbreeding and recombination between populations. Speciation encompasses the diversification of all forms of phenotype, including the evolution of phenotypic differences that make possible the exploitation of different ecological resources, and therefore allowing individuals to coexist in sympatry (Barton, 2001).

The concept of 'sibling species' (Mayr, 1963) has received considerable attention in evolutionary studies of fishes, especially species found in north temperate lakes (Taylor, 1999). Sibling species are populations that appear morphologically similar, but differ in ecology or life history (e.g. different maturity size, spawning in different areas of the same lake at different

times) and display some degree of genetic differentiation, suggestive of reproductive isolation (Mayr, 1963; Taylor, 1999). Morphological analyses of these sibling species revealed that most of them showed some phenotypic divergence (Taylor, 1999). Biochemical and molecular assays have also revealed varying levels of genetic differentiation, likely due to behavioural and/or morphological differences of these forms existing sympatrically (Taylor, 1999). Such studies of morphology and molecular genetics eventually led to the discovery of many cases of sympatric pairs of fishes (Taylor, 1999). These are species that occur in sympatry, for example in the same lake or river system, and are identified by a common Latin binomial but are reproductively isolated from one another (Taylor, 1999).

Many cases of sympatric species pairs have been documented in freshwater ecosystems of post-glacial lakes in the Northern Hemisphere. These sympatric pairs are often used as model systems to study the early stages of population divergence and adaptation, ultimately leading to speciation (Saint- Laurent *et al*., 2003). Some of the most studied families that include sympatric pairs in post glacial lakes are Salmonidae (trout, salmon, grayling whitefish and charr), Gasterosteidae (sticklebacks) and Osmeridae (smelts) (Taylor, 1999). While many of these cases are in post-glacial freshwater ecosystems in the temperate regions of the Northern Hemisphere they are not exclusive to those areas (Taylor, 1999). Other documented cases include ecologically and genetically isolated forms of the loach (*Cobitistaenia*, Cobiditae) in Honshu, Japan (Minamori, 1956). An allozyme survey conducted on goodeiid fish *Ilyodon,* Goodeidae (Turner & Grosse, 1980) described ecotypes that were reproductively isolated in at least one site in Mexico. There is also evidence of sympatric pairs of *Percichthystrucha* (Pisces: Percichthyidae) found in six lakes in the Andean region of Argentina (Ruzzante *et al*., 1998). Gill raker length and morphological measures were all associated with the use of different

habitat use and diet. Finally, there are also records of a putative sympatric species pair in the marine fish *Acanthochromis polyacanthus*, Pomacentridae (damselfish) and *Ammodytes*, Ammodytidae (sand lances) (Okamoto, 1989; Planes & Doherty, 1997). All of these sympatric pairs differ in one way or another either due to their trophic ecology (e.g. benthic, planktivorous, etc.), life history (such as size at maturity) (Behnke, 1972), morphology and behavior (Taylor & Bentzen, 1993a; Taylor, 1999). Studies have also shown that the genetics of these sympatric populations of fish differ from one another, which suggests reproductive isolation (Taylor & Bentzen, 1993).

It is believed that these sympatric populations have evolved via ecological speciation. Ecological speciation is a theory that joins speciation processes by which reproductive isolation arises as a result of disruptive or divergent selection on traits in different environments (Schluter, 2001). These different environments can be natural or even unnatural elements of their habitat, such as food resources, climate and even physical structures or barriers (Schluter, 2001). Interactions with other species, such as mutualism or predation and resource competition can also all have some influence on selection (Schluter, 2001).

The underlying concept is that the depletion of shared resources by species that are closely related and morphologically similar supports phenotypes that exploit new resources; groups of individuals are subjected to differing selection pressures leading to phenotypic divergences (Schluter, 2000). These genetic variations however, evolve using alleles that were already present within the common ancestral population, and not due to new mutations (Schluter & Conte, 2009). Reproductive isolation then increases between populations as a by-product of adaptation to different selection regimes (Schluter, 2001). The phenotypic differences observed in sympatric populations are often associated with different ecological or

trophic strategies, as documented in freshwater pairs of Threespine Stickleback (*Gasterosteus aculeatus* (Schluter & McPhail, 1992; Taylor, 1999), Lake Whitefish (*Coregonus clupeaformis*) (Lu & Bernatchez, 1999) and Arctic Charr (*Salvelinus alpinus*) (Snorrason *et al.* 1994).

Resource polymorphisms occur quite often in vertebrate species, (Skúlason & Smith, 1995; Smith and Skúlason, 1996). Resource partitioning appears to be promoted by two conditions: the existence of empty niches or underused resources, and a related relaxation of competition (Smith and Skulason, 1996). This situation is common in post-glacial lakes and is consistent with the high occurrence of species, or population pairs exploiting alternative resources in their biotic community (Skúlason & Smith, 1995; Smith and Skúlason, 1996; Taylor & McPhail, 1999; Robinson & Schluter, 2000).

The most prominent morphological differences between sympatric forms are associated with preferred trophic niche. Examples include benthic and limnetic sticklebacks (McPhail, 1994; Taylor, 1999), planktivorous and epibenthivorous whitefish (Lu & Bernatchez, 1999), planktivorous, benthivorous and piscivorous Arctic Charr (Snorrason *et al.* 1994), and microphagous and macrophagous forms of smelt (Taylor & Bentzen, 1993b). Benthic and epibenthic stickleback, whitefish and charr as well as the piscivorous macrophagous smelt have fewer gill rakers than limnetics, planktivorous and microphagous fish (McPhail, 1984; Baby *et al*., 1991; Taylor & Bentzen, 1993b; Snorrason *et al.* 1994; Lu & Bernatchez, 1999). Elevated numbers of gill rakers allows for more efficient selection of small prey such as zooplankton (Lu & Bernatchez, 1999), whereas benthivorous fish have been found to consume larger prey and thus have fewer gill rakers (McPhail, 1984).

Such adaptive radiations are common following the recolonization of previously glaciated habitats (Taylor & Bentzen, 1993b; Bradbury *et al*., 2010), and have led to morphologically bimodal populations, in which large differences in phenotype exist (Bradbury *et al*., 2010). Because these adaptive radiations in sympatric pairs are recent (<10 000 – 15 000 years), they have the potential to provide information on early speciation in action (Bradbury *et al.*, 2010).

1.2 Study Species

The Rainbow Smelt (*Osmerus mordax*, Mitchill) is a small osmerid fish found in freshwater and coastal systems ranging from the Delaware River, USA to Labrador, and the Gulf of St. Lawrence Canada (Nellbring, 1989; Buckley, 1989; Bradbury *et al.* 2010). Smelt also occur naturally in ponds and lakes in northeastern North America (Nellbring, 1989). In the early 1900's they were introduced to and are now abundant in the Great Lakes (Nellbring, 1989). Along with their vast distribution, Rainbow Smelt display widespread phenotypic and life history diversity (Taylor & Bentzen, 1993). Two major life history types exist: sea-run (anadromous), and permanent lake resident (lacustrine) populations (Lanteigne & McAllister, 1983; Taylor & Bentzen, 1993a&b; Bernatchez, 1997). Anadromous smelt spawn in freshwater streams and grow and mature in estuaries and coastal waters (Buckley, 1989). Freshwater smelt have similar spawning patterns, in which they migrate into streams to spawn (Buckley, 1989). Spawning typically occurs from late March to late May, after dark at night, on a gravel substrate, in water depths of 0.1- 1.3 m, (Murawski *et al.*, 1980; Buckley, 1989) with preference given to high velocity water (Buckley, 1989). Smelt are broadcast spawners; eggs and milt are released in no particular pattern into their spawning streams or rivers. On

any given night during the spawning run, male smelt predominate, which can be attributed to the fact that males have a longer spawning period (Buckley, 1989). Rupp (1968) has reported that an individual male smelt can spawn eight nights in a row, whereas females are found to only spawn three to four consecutive nights.

Sympatric forms of Rainbow Smelt have been observed in many lakes and estuaries in northeastern North America. They have been identified in Heney Lake, Quebec, Lake Utopia New Brunswick, Green Lake and Wilton Pont in Maine, USA (Lanteigne & McAllister, 1983), Lochaber Lake, Nova Scotia (Taylor & Bentzen, 1993b), the St. Lawrence estuary in Quebec (Bernatchez & Martin, 1996) and Lac Saint-Jean, Quebec (Saint-Laurent *et al*., 2003).

The sympatric forms of Rainbow Smelt in Lochaber Lake are diversified into microphagous and macrophagous forms, where the microphagous smelt mature at 105-155mm fork length, have larger eyes, shorter upper jaws and more gill rakers than the macrophagous smelt which mature at 185-300mm fork length (Taylor & Bentzen, 1993b). Both forms spawn in the same two streams (McNab and Hurlbert Brooks), and even though there may be a small difference in peak spawning date, both forms are found spawning in the same streams on the same nights (Taylor & Bentzen, 1993b; P. Bentzen pers. com.).

Previous phylogeographic studies of mtDNA showed that there are two distinct lineages of smelt that evolved in allopatry for hundred thousands of years (Taylor & Bentzen, 1993b; Bernatchez, 1997). Secondary contact between the two lineages is one potential explanation for morphological and genetic differences observed in landlocked populations of smelt. Phylogeographic studies using mtDNA of Lake Whitefish sympatric pairs revealed that these ecotypes result from secondary contact between different glacial races that have

been isolated from one another since the Illinoian glacial period (150 000 years ago) (Bernatchez & Dodson, 1990, 1991). The dwarf ecotype typically associates with the Acadian glacial race, whereas the normal ecotype is associated with either the Atlantic glacial race or the Mississippian glacial race, depending on the lake (Bernatchez & Dodson, 1991). Taylor & Bentzen (1993b) showed however that the sympatric forms of smelt in Lochaber Lake and Lake Utopia are derived from a single glacial lineage, and that the two forms have likely diverged independently in each lake. More broadly, Taylor and Bentzen (1993b) showed that small microphagous and large macrophagous freshwater smelt are polyphyletic in origin, that is, they have evolved independently in a number of lakes, presumably from anadromous ancestors.

1.3 Thesis Objectives

This thesis uses molecular genetic methods to examine population structure of small and large morph Rainbow Smelt in Lochaber Lake. Patterns of population structure will be used to clarify the divergence between the small and the large morphs of smelt. Samples included in this study were collected in the two spawning streams and in lake's water column.

Chapter 2 examines patterns of genetic differentiation of spawning small and large morph smelt in Lochaber Lake. Smelt samples collected in five non-consecutive years over a 10-year period were genotyped for microsatellite loci. Objectives of this research were to (1) quantify genetic differentiation between the two morphs, (2) estimate hybridization rates and (3) assess the stability of these measures over time.

Chapter 3 examines genetic relationships between the adults of both morphs and larval smelt collected over a two-year period. Objectives this research were to (1)

determine whether larvae could be identified to morph using microsatellite markers and (2) determine whether hybrid status (e.g. pure vs. hybrid) could be ascertained, and (3) if so, to test whether the frequency of hybrids differs between larvae and adult spawners.

CHAPTER 2 Genetic differentiation of the small and large Rainbow smelt (*Osmerus mordax*) in Lochaber Lake, Nova Scotia and its stability over time

2.1 Intoduction

Explaining reasons for population divergence, and speciation, remain important objectives in evolutionary biology (Schluter, 2000; Barton, 2001). Speciation is one of the central features of evolution, yet it still remains incompletely understood (Schluter, 2001). The role of natural selection as well as neutral evolutionary processes in producing and maintaining biological diversity has long been debated. There are two mechanisms of speciation by selection: mutation-order and ecological speciation (Schluter, 2010). Speciation under mutation- order occurs when many different mutations arise and get fixed in populations that are adapting to similar environments; whereas, in ecological speciation differences are driven by natural selection in different environments (Schluter, 2010). When referring to ecological speciation, we refer to the evolution of reproductive isolation between populations or a subset of a single population due to adaptations to different niches or environments (Schluter, 2010; 2001; Rundle & Nosil, 2005). For sympatric populations, resource partitioning encompasses the exploitation of different niches to avoid or reduce competition through the evolution of resource-based polymorphisms (Skúlason & Smith, 1995). According to the ecological speciation hypothesis (Schluter, 1996a&b; 2001) the same selective pressures leading to resource-based trophic polymorphism may also favour reproductive isolation through reinforcement.

Adaptive radiations following recolonization of previously glaciated habitats have been reported in a number of temperate fishes including freshwater pairs of stickleback, lake

whitefish and Arctic charr (Claytor & Verspoor, 1991; Taylor & Bentzen, 1993b; Bernatchez, Chouinard & Lu, 1999; Blackie, Weese & Noaks, 2003; Bradbury et al., 2010, Osinovet al., 2015; Harris et al., 2015). These recent radiations have often resulted in morphologically bimodal populations with multiple phenotypic differences (Bell, 1976, 1989; Lanteigne & McAllister, 1983; Bradbury et al., 2010). Because these adaptive radiations in sympatric pairs are recent (<10 000-15 000 years, since the last glaciation), they have the potential to provide information on early speciation in action (Bradbury et al., 2010).

In many cases of post-glacially derived sympatric species pairs of fishes, morphological differences are likely associated with preferred habitat and trophic niches. In sticklebacks, it has been observed that there are benthic and limnetic populations that coexist in up to seven lakes in British Columbia, Canada (Schluter & McPhail, 1992; McPhail, 1994; Taylor, 1999). There are also 'dwarf' (planktivorous) and 'normal' (epibenthivorous) Lake Whitefish (Fenderson 1964; Chouinard et al., 1996; Lu & Bernatchez, 1999) and planktivorous, benthivorous and piscivorous Arctic Charr (Snorrason et al., 1994). In all these cases, the benthic stickleback, epibenthic and piscivorous whitefish and charr, all have larger bodies, lower gill raker counts and larger mouths than the limnetic and planktivorous forms (McPhail, 1984; Lu & Bernatchez, 1999; Snorrason *et al*., 1994; Baby et al., 1991). Elevated numbers of gill rakers in the smaller forms allow for efficient selection of small prey (Lu & Bernatchez, 1999), whereas benthivorous fish likely consume larger prey and thus have lower gill raker counts as they do not need to filter for small prey (McPhail, 1984).

Lacustrine Rainbow Smelt are found in numerous lakes in Atlantic Canada. In at least two Maritime lakes, Lake Utopia (New Brunswick) and Lochaber Lake (Nova Scotia), smelt are further diversified in sympatric populations of large bodied and small bodied forms. In both

lakes, the small bodied smelt mature at 105-155mm fork length, have larger eyes, shorter upper jaws and more gill rakers than large bodied smelt which mature at 185-300mm fork length (Lanteigne & McAllister, 1983; Taylor & Bentzen, 1993a&b). Several studies of sympatric fish pairs have demonstrated that these morphological differences are primarily associated with trophic preferences (Taylor & Bentzen, 1993a&b). It is therefore assumed that trophic ecology plays a very important role in smelt life history differentiation (Taylor & Bentzen, 1993b), whereby the small morph is thought to be primarily planktivorous, (Copeman, 1977; Copeman & McAllister, 1978) and the large morph is believed to be piscivorous.

Fork lengths (FL) were available for adults caught in Lochaber Lake in 2004 and 2005 (Figure 2.1). In both years the length frequency distributions of the small and large morphs showed no overlap. Small morph smelt had fork lengths of 105-147mm for 2004 (mean = 123.8mm) and 107-155mm FL in 2005 (mean = 127mm), and large morph smelt had fork lengths of 186-299mm FL for 2004 (mean = 243.4mm) and 185-300mm FL for 2005 (mean = 231.7mm). Size length differences between the small and the large morphs appear to be persistent over time and that the small and the large morph are easily distinguishable. The persistent morphological differences between the small bodied and large bodied smelt are thought to be the product of disruptive selection (Bradbury *et al*., 2010), which is when reproductive isolation between populations or subsets of a single population develops as a result of adaptations to dissimilar ecological niches or environments and resources (Hatfield & Schluter, 1999; Schluter, 2009). Based on mitochondrial DNA (mtDNA) analyses, Taylor & Bentzen (1993b) documented that the two smelt morphs in Lochaber are more closely related to each other than either is to the corresponding morph in Lake Utopia, implying that freshwater smelt morphs have arisen independently in different lakes, and thus are polyphyletic in origin.

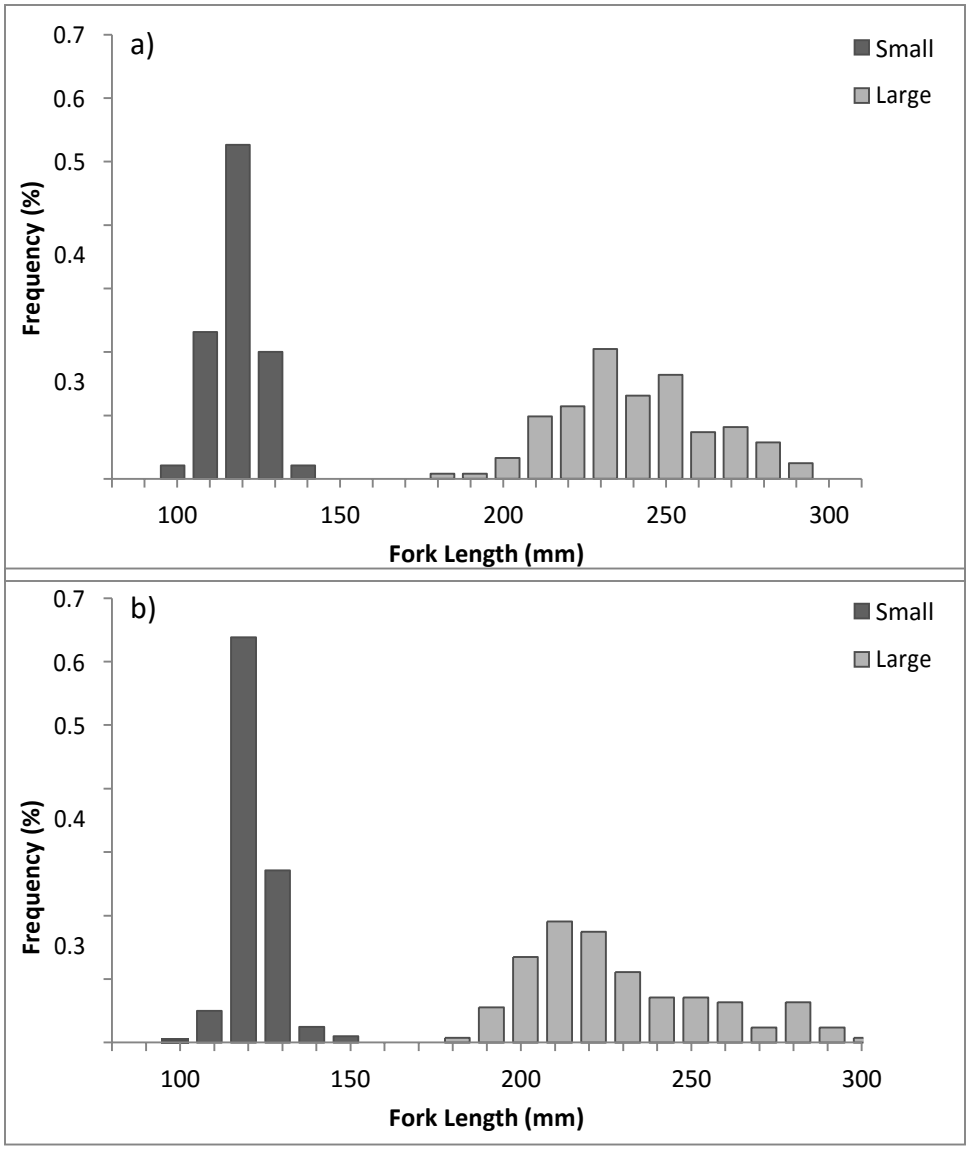


Figure 2.1. Length frequency distributions of spawning smelt in Lochaber Lake for smelt caught in 2004 (a) where small morphs measured between 105 and 147mm FL, and the large morphs measure between 186 and 299mm FL and 2005 (b) where small morph measured between 107 and 155mm FL and large morph measured between 185 and 300mm FL. Note the lack of overlap in length distribution.

Both forms of smelt in Lochaber Lake spawn in the same two streams (McNab and Hurlbert Brooks) that flow into the lake. Although the possibility that peak spawning dates differ cannot be ruled out, both forms are found spawning in the same location in both streams on the same nights (Taylor & Bentzen, 1993b; P. Bentzen pers. obs.). Because of the temporal overlap, and the fact that both forms are broadcast spawners, there is potential for gene flow between the two ecotypes. Bradbury *et al.* (2010) conducted laboratory fertilization experiments on the two ecotypes from Lochaber Lake that revealed that fertilization between morphs was possible, although with a 30-50% decrease in successful fertilization relative to the fertilization rate within morphs (~78%).

Taylor and Bentzen (1993b), reported differentiation between the two morphs in Lochaber Lake, but these results were subject to certain limitations. They only examined a small sample of each morph (N= 54; 37 small and 17 large morph), and given that mtDNA is maternally inherited, it represents only a small fraction of the total historical record in a sexual organismal pedigree (Palumbi & Baker, 1994). Furthermore, mtDNA has a fourfold smaller effective population size (N_e) than nuclear markers, which can enhance the effects of genetic drift in populations that are subdivided, and may lead to faster fixation or loss of alleles and stronger populations subdivisions at mitochondrial compared to nuclear loci (Birky *et al*., 1983). While the overall mutation rates tend to be greater for the mitochondrial genome than for the nuclear genome (Brown *et al*., 1979), much of the mitochondrial genome is protein coding and therefore possibly under selection (Ballard & Kreitman, 1995), meaning it may not always evolve quickly enough to infer levels of contemporary gene flow (Angers & Bernatchez, 1998). In part because of these limitations of mtDNA, microsatellites have become the most commonly used nuclear markers (Estoup & Angers, 1998) because microsatellites are: 1) inherited from

both parents, which means that most loci are selectively neutral and mutations accumulate rapidly (Balloux & Lugon-Moulin, 2002) and 2) polymorphic in nature and therefore typically have high information content which provide better discriminatory power for resolving population structure than mtDNA (Goudet *et al*., 1996).

2.1.1 Chapter Objectives

In the study conducted by Bradbury *et al*. (2010), microsatellite markers were used to study genetic differentiation in Lochaber Lake smelt, but data were limited to 50 individuals of each morph that were collected over two years only. This chapter examines patterns of genetic differentiation of adult spawning small and large morph smelt in Lochaber Lake. Smelt samples collected in five non-consecutive years over a 10 year period were genotyped at multiple microsatellite loci. Objectives of this research were to quantify genetic differentiation between the two morphs, estimate the rate of hybridization and assess the stability of these measures over time.

2.2 Materials and Methods

2.2.1 Study Area

Lochaber Lake (45º25', 62º01') is situated in northeastern mainland Nova Scotia, Canada straddling the border between Antigonish and Guysborough Counties (Figure 2.2). Lochaber Lake is a small but deep lake with a surface area of 3km^2, a maximum width of 0.75km and a length of 4.5km. The average depth of the lake is 21.8m with a maximum depth of 52.4m near the center of the lake. Lochaber Lake drains into the St. Mary's River system, flowing south and eventually draining into the Atlantic Ocean at Sherbrooke, NS. There are eight streams that feed the lake on a nearly continuous basis. The fish community of the lake includes White Sucker

(*Catostomus commersoni*), American Eel (*Anguilla rostrata*), Yellow Perch (*Perca flavescens*), White Perch (*Morone americana*), Brook Trout (*Salvelinus fontinalis*), Brown Bullhead (*Ameiurus nebulosus*), Threespine Stickleback (*Gasterosteus aculeatus*), Ninespine Stickleback (*Pungitius pungitius*), Banded Killifish (*Fundulus diaphanous*) and Rainbow Smelt (*Osmerus mordax*).

There are two tributaries in Lochaber Lake known to support smelt spawning, McNab and Hurlbert Brooks (Figure 2.2). Both smelt ecotypes spawn in both streams at the same time (late March to early or mid-April depending on ice melt) although peak spawning time may differ slightly between ecotypes. Both streams are shallow, and have a rocky substrate with little to no vegetation. Hurlbert Brook has a raised culvert approximately 130m from the lake limiting smelt migration further upstream (Figure 2.2), whereas McNab Brook is long and narrow with no man-made barrier, although no egg deposition was found at distances greater than approximately 200m upstream.

2.2.2 Field Methods

Adult smelt of both morphs were collected during spawning periods in April 2002, 2003 and 2004, and in March and April of 2010 and 2011 in McNab and Hurlbert Brooks. All fish were caught after dark with the use of small dip nets and kept in large buckets filled with stream water for a few minutes prior to sampling. The ecotype, fork length (for the 2004 & 2005 samples only) and sex of each fish was noted and a small pectoral or caudal fin clip was taken from every smelt and preserved in tubes containing 100% ethanol until further use. Each clip was given a corresponding number with information regarding the ecotype, length and the sex of the particular fish. A total of 675 small and 236 large morph smelt were included in the dataset.

2.2.3 DNA isolation and data collection

Data for the 2002-2004 smelt samples (517 small and 197 large) were previously collected by a laboratory technician using the same protocols employed for contemporary samples. Fin tissue from 159 small and 38 large adult smelts collected in 2010 and 2011 were placed in 225µL of digestion buffer (100mM NaCl, 50mM TrisHCl pH8, 10mM EDTA, 0.5% SDS) with 2µL proteinase K and incubated overnight at 55ºC and 200rpm in an orbital shaker. DNA was extracted using the glass milk protocol described by Elphinstone *et al.* (2003), and modified to work with a 96-well filter plate and automated using a Perkin Elmer MultiPROBEII liquid handling system. Extracted DNA was stored at -20ºC in preparation for further analysis. Nine neutral polymorphic microsatellite loci developed for *O. mordax* were used as follows for the 2002-2004 samples: *Omo1, Omo2, Omo3, Omo4, Omo5, Omo6, Omo11, Omo14& Omo16* (Coulson *et al.*, 2006). Ten microsatellite loci were used for the 2010 and 2011 samples as follows: *Omo1, Omo2, Omo3, Omo4, Omo5, Omo9, Omo11, Omo13, Omo14& Omo15.* The two data sets shared seven loci (*Omo1, Omo2, Omo3, Omo4, Omo5, Omo11 & Omo14*), which were used for this study. Allele frequency distributions for the 2002-2004 data and 2010-2011 data were compared to ensure standardization of alleles across the combined data set. Individuals were genotyped using polymerase chain reaction (PCR) in 5- or 10- µL volumes containing 20-100ng DNA, 1.5mM $MgCl_2$, 80µM of each dNTP, 0.5 U *Taq* DNA polymerase (Applied Biosystems), 0.3 µM of each primer if the forward primers were end labeled with HEX dye or 0.01µM of forward primer, 1µM of reverse and 1µM of a HEX labeled tailed primer and 1x PCR buffer (10mM Tris-HCl, pH8.3; 50mM KCl). Two temperature profiles were used for touchdown PCRs.

Touchdown PCR was used for all primer pairs: OmoTD2 program was used for *Omo1, Omo2, Omo3, Omo4* and *Omo5*, and OmoTD5 was used for *Omo11* and *Omo14.*

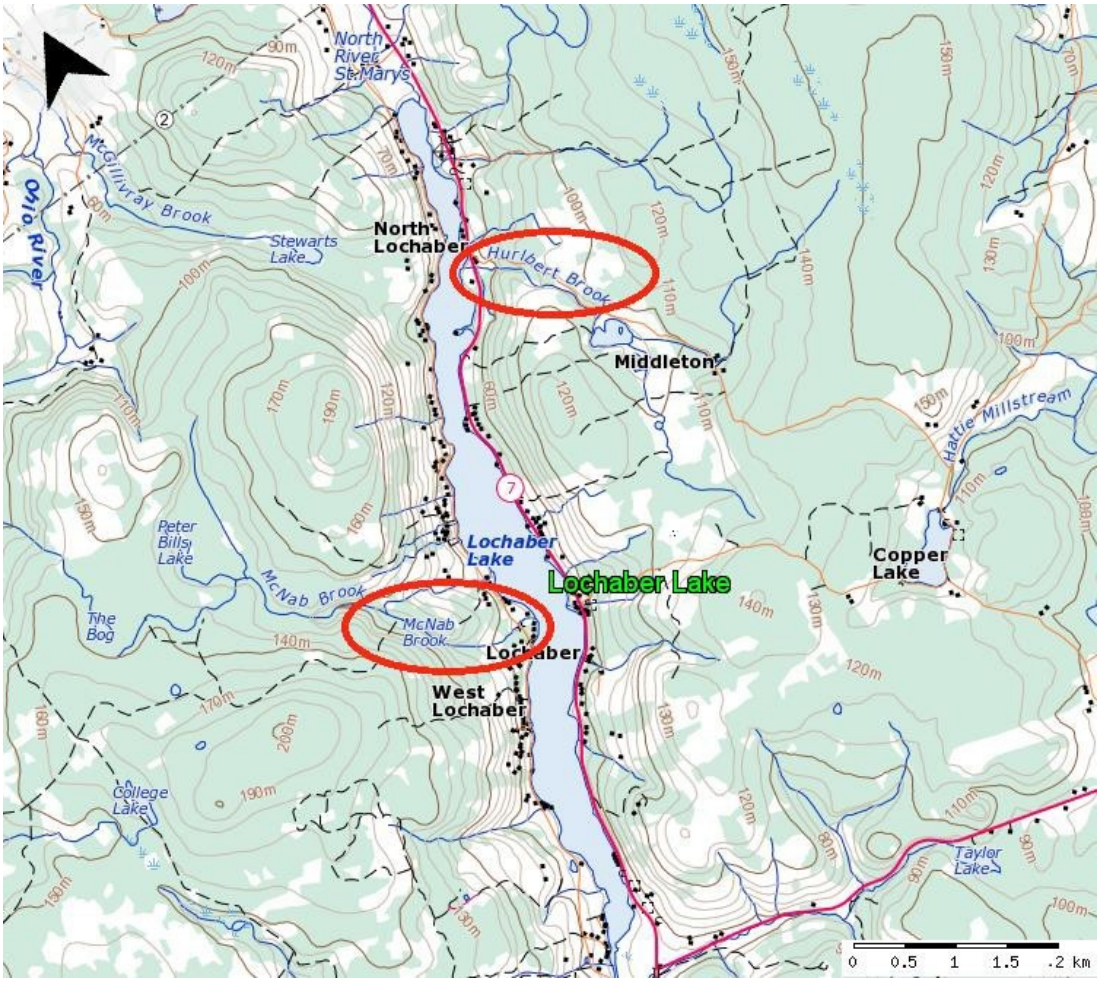


Figure 2.2: A topographical map taken from Natural Resources Canada shows Lochaber Lake (N45°25'1", W62°1'41"), where both Hurlbert and McNab brooks are circled in red. Many other streams flow into or out of Lochaber, however no evidence of smelt spawning has been found in them.

Touchdown PCR conditions were as follows for OmoTD2: 94ºC for 2min, followed by five cycles of 94ºC for 30 s, 66ºC for 30 s (-1ºC for the following four cycles), 72ºC for 30 s, followed by 25 cycles of 94ºC for 30 s, 62ºC for 30 s and 72ºC for 30 s with a final extension was held at 72ºC for 5 min. For OmoTD5, PCR conditions were as follows: 94ºC for 2min, followed by four cycles of 94ºC for 30 s, 64ºC for 30 s (-1ºC for the following three cycles), 72ºC for 30 s, followed by 26 cycles of 94ºC for 30 s, 60ºC for 30 s and 72ºC for 30 s with a final extension was held at 72ºC for 5 min. Reactions were run on Eppendorf thermocyclers and visualized using a denaturing polyacrylamide gel electrophoresis on an FMBioII system (Hitachi Genetic Systems). Genotypes were scored by hand in reference to a molecular weight size standard ladder incorporated into each gel. Every image also incorporated a positive control and a redundant sample in order to allow comparison of new samples to those of known allele sizes which helped ensure consistent scoring across both ecotypes and loci.

2.2.4 Data Analyses

Microsatellite scores were checked for genotyping errors, such as null alleles, stutter and large allele drop out were not present using MICRO-CHECKER v2.2.3 (van Oosterhout *et al*., 2004). For each locus and both ecotypes the total number of alleles, the size range of alleles, expected- (H_E) and observed (H_O) heterozygosity, and inbreeding coefficient (F_{IS}) were calculated using GenAlEx v.6.41 (Peakall & Smouse, 2006); allelic richness (A_E) for all loci and both morphs, as well as the average microsatellite heterozygosity for each morph were calculated using FSTAT v. 1.2 (Goudet, 1995). Deviations from HWE were assessed using the Markov chain Monte Carlo (MCMC) approximation of Fisher's exact test and linkage disequilibrium (LD) was assessed for all possible loci combinations using the

MCMC simulated exact tests as implemented in ARLEQUIN 3.5 (Excoffier & Lischer, 2010). These results are reported in Appendix 2.1.

In tests for both HWE and LD, statistical significance (α) was adjusted for the number of simultaneous tests *k* (α/k for $\alpha = 0.05$) using a sequential Bonferroni procedure (Rice, 1989) in order to reduce Type I errors. Because the action of selection can generate patterns of LD and violate some of the critical assumptions of HWE, both ecotypes were evaluated separately. Loci were removed only if LD or deviation from HWE was detected for both ecotypes.

Of the seven loci, Omo14 showed deviations from HWE for both ecotypes therefore all of the following analyses were done with and without this locus. A second locus, Omo2, was not used in the collection of data for two of the five sampling years, so analyses were also done with and without the presence of this locus. Analyses were also performed with all seven loci, and on a data set where both Omo2 and Omo14 were removed, leaving five loci in the data set.

Estimates of pair-wise F_{ST} between the two ecotypes, and their statistical significance, were calculated using Microsatellite Analyser (MSA) (Dieringer & Schlötterer, 2003). Principal Coordinate Analysis (PCoA) were carried out using a distance matrix created from the codominant genotypic distance algorithm of GenAlEx v.6.41 (Peakall & Smouse, 2006, algorithm described in Smouse & Peakall, 1999). PCoA's (also generated using GenAlEx) were done by year as well as for the entire data set.

Bayesian clustering analysis was conducted in STRUCTURE v.2.3.3 (Pritchard *et al*., 2000) and tested the presence of multiple discrete populations. This method uses assumptions of linkage equilibrium and HWE among loci, establishes population structure, and uses a

MCMC algorithm to identify populations that are not in linkage disequilibrium and develop estimates of the number of populations (K). The program was run four times for each value of K to assess the stability of values, and with a burn-in of 100,000 iterations and 500,000 iterations after burn-in, and K ranging from 1 to 5. The most likely K-value was inferred based on the rate of change of the log probability of data (ΔK) (Evanno *et al*., 2005) using STRUCTURE HARVESTER (Earl & von Holdt, 2011). STRUCTURE also estimates q-values, or admixture coefficients, for each individual, which represents the portion of each multilocus genotype allocated to each cluster or population. If hybridization occurs between the two ecotypes, hybrids should possess an intermediate q-value. Results from replicate STRUCTURE runs were merged using CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) and graphically summarized using DISTRUCT 1.1 (Rosenberg 2004).

Migration rates between the two morphs were estimated using a Bayesian approach implemented in BAYESASS v.3.0.1 (Wilson & Rannala, 2003). This was done for each year separately, and for the entire data set combined. BAYESASS was run with a burn-in of 1 million MCMC iterations followed by 10 million sample iterations. This process was done several times for each data set, each time with a different starting seed number to ensure consistent and accurate results, as recommended in the program manual. The results from all the runs were examined to make sure that there was convergence (by comparing the similarity of the results), but only the results from the first BAYESASS run were reported. Standard deviations given by the program were used to construct 95% confidence intervals around each estimated rate of migration. If the interval did not include zero, migration rates were considered significant.

2.3 Results

2.3.1 Data quality

Of the seven loci used in the data set, MICROCHECKER found that for the 2002-2004 data, Omo2 & Omo14 had null alleles in the small morph, and Omo4 and Omo14 appeared to have null alleles in the large morph. For the 2010-2011 samples, the small smelt population showed no evidence of null alleles, whereas the large smelt appeared to have null alleles at Omo11.

No locus pair consistently showed patterns of LD; however, in the 2004 small morph data set Omo1 & Omo14 appeared to be in linkage disequilibrium. One locus, Omo14, did not appear to be in HWE in both ecotypes and in all years sampled. Omo11 deviated from HWE in small morphs collected in 2004 and large morphs collected in 2010/2011. Two loci, Omo3 and Omo4 also appeared to deviate from HWE in the large morphs collected in 2004 (Appendix 2.2). Because these patterns were not consistent in both morphs in all years, all loci were kept in the analyses. Analyses were done four times, each with different loci combination in order to compare results. All seven loci were used and compared to analyses where (1) only Omo2 was removed, (2) only Omo14 was removed and where (3) both Omo2 and Omo14 were removed.

2.3.2 Genetic Analysis

Average microsatellite heterozygosities for the combined data set were lower in the large morph (0.57) than in the small morph (0.75), but this difference was not significant (ANOVA, $p=0.068$). The average number of alleles did significantly differ (ANOVA, $p=0.044$) between the large ($N_a=7.14$) and the small ($N_a=12.43$) morphs. Pairwise F_{ST} values between morphs for

smelt ranged from 0.070 to 0.116 (Table 2.1). In general, removing the two problematic loci (Omo2 and Omo14) increased the F_{ST} for the combined data set and for each year separately. All F_{ST} values differed significantly from zero ($p<0.05$), indicating significant genetic differentiation between the two morphs. Genetic distances among individuals were examined using Principle Coordinate Analysis (PCoA) for all years and combinations of loci. These analyses revealed differences between the morphs but also showed some overlap (Figures 2.3-2.6). In 2002 (figures 2.3a-2.3d) and 2003 (figures 2.4a-2.4d), results were very similar for all combinations of loci. All PCoA figures showed differences between the small and the large smelt, although some overlap was evident where the more variable small smelt appear to overlap the more constricted genotypic distribution of the large morph. The small smelt also appeared more genetically variable than the large smelt. For the 2004 smelt (figures 2.5a-2.5d) and the 2010-2011 smelt (figures 2.6a-2.6d) the same pattern was seen but the overlap was not as apparent. When the entire data set was combined, the two morphs of smelt still showed independence from one another but overlap was still observed (figures 2.7a and 2.7b). Note that Omo2 was not included in any PCoA plots for the combined data as results were skewed due to it not having been genotyped for the 2002 and 2003 samples.

Bayesian clustering of the data for all years using STRUCTURE supported the presence of two genetic groups with K = 2 being the most likely number of clusters determined by the tests implemented in STRUCTURE HARVESTER (Appendix 3). The two clusters correspond to the small and the large morphs of smelt sampled from both streams in the lake (Figure 2.8). The two clusters correspond to the small and the large morphs of smelt sampled from both streams in the lake (Figure 2.8).

Table 2.1. Pairwise F_{ST} values and p-values for all samples of rainbow smelt from Lochaber Lake calculated using a permutation test implemented in MSA.

Loci Combination	Year →	2002		2003		2004		2010/2011		Combined	
	Ecotype ↓	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
All Loci	Small	---	0.103	---	0.087	---	0.075	---	0.074	---	0.074
	Large	<0.001	---	<0.001	---	<0.001	---	<0.001	---	<0.001	---
Omo2 Removed	Small	---	0.104	---	0.087	---	0.081	---	0.076	---	0.070
	Large	<0.001	---	<0.001	---	<0.001	---	<0.001	---	<0.001	---
Omo14 Removed	Small	---	0.116	---	0.099	---	0.071	---	0.077	---	0.074
	Large	<0.001	---	<0.001	---	<0.001	---	<0.001	---	<0.001	---
Omo2&14 Removed	Small	---	0.116	---	0.099	---	0.078	---	0.081	---	0.081
	Large	<0.001	---	<0.001	---	<0.001	---	<0.001	---	<0.001	---

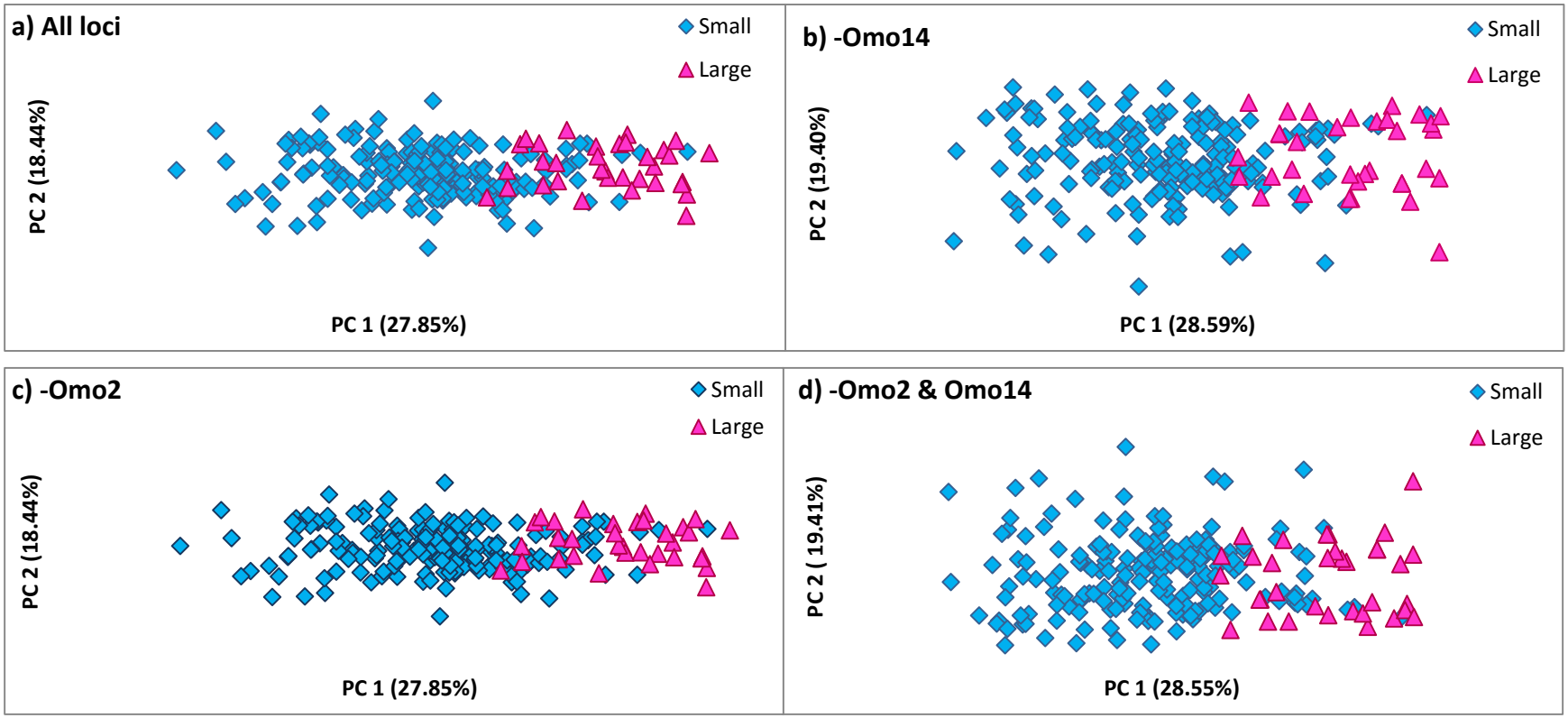


Figure 2.3. Principle coordinate analysis for adult smelt ecotypes collected in 2002 (N=225, 191 small & 34 large morph). Results shown are a) all 7 microsatellite loci; b) 6 microsatellite loci with Omo14 removed; c) 6 microsatellite loci with Omo2 removed and d) 5 microsatellite loci with both Omo2 & Omo14 removed. All four plots show two dimensions of genetic variation; numbers in parentheses show percentage of genetic variation explained by one axis.

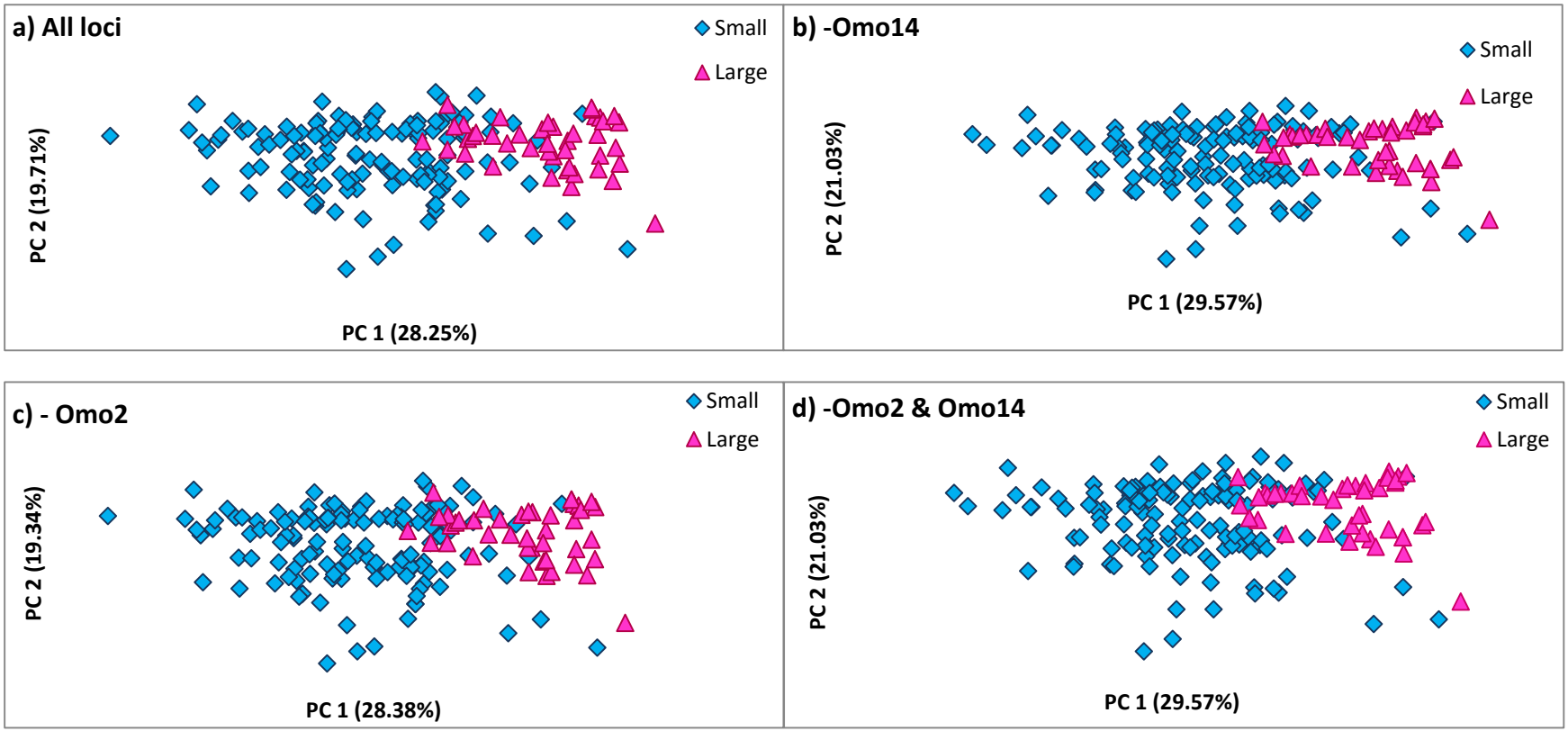


Figure 2.4. Principle coordinate analysis for adult smelt ecotypes collected in 2003 (N=179, 136 small & 43 large morph). Results shown are a) all 7 microsatellite loci; b) 6 microsatellite loci with Omo14 removed; c) 6 microsatellite loci with Omo2 removed and d) 5 microsatellite loci with both Omo2 & Omo14 removed. All four plots show two dimensions of genetic variation; numbers in parentheses show percentage of genetic variation explained by one axis.

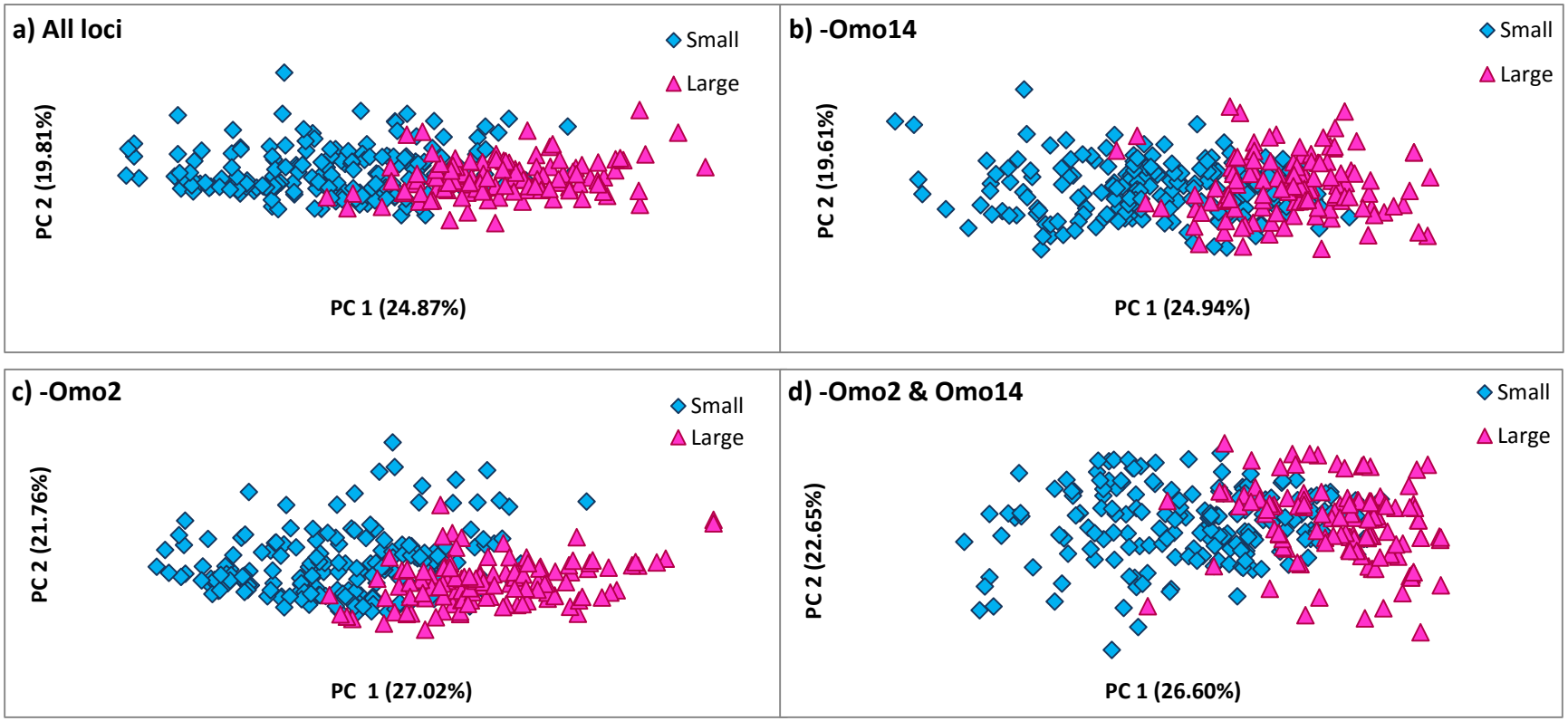


Figure 2.5. Principle coordinate analysis for adult smelt ecotypes collected in 2004 (N=310, 190 small & 120 large morph). Results shown are a) all 7 microsatellite loci; b) 6 microsatellite loci with Omo14 removed; c) 6 microsatellite loci with Omo2 removed and d) 5 microsatellite loci with both Omo2 & Omo14 removed. All four plots show two dimensions of genetic variation; numbers in parentheses show percentage of genetic variation explained by one axis.

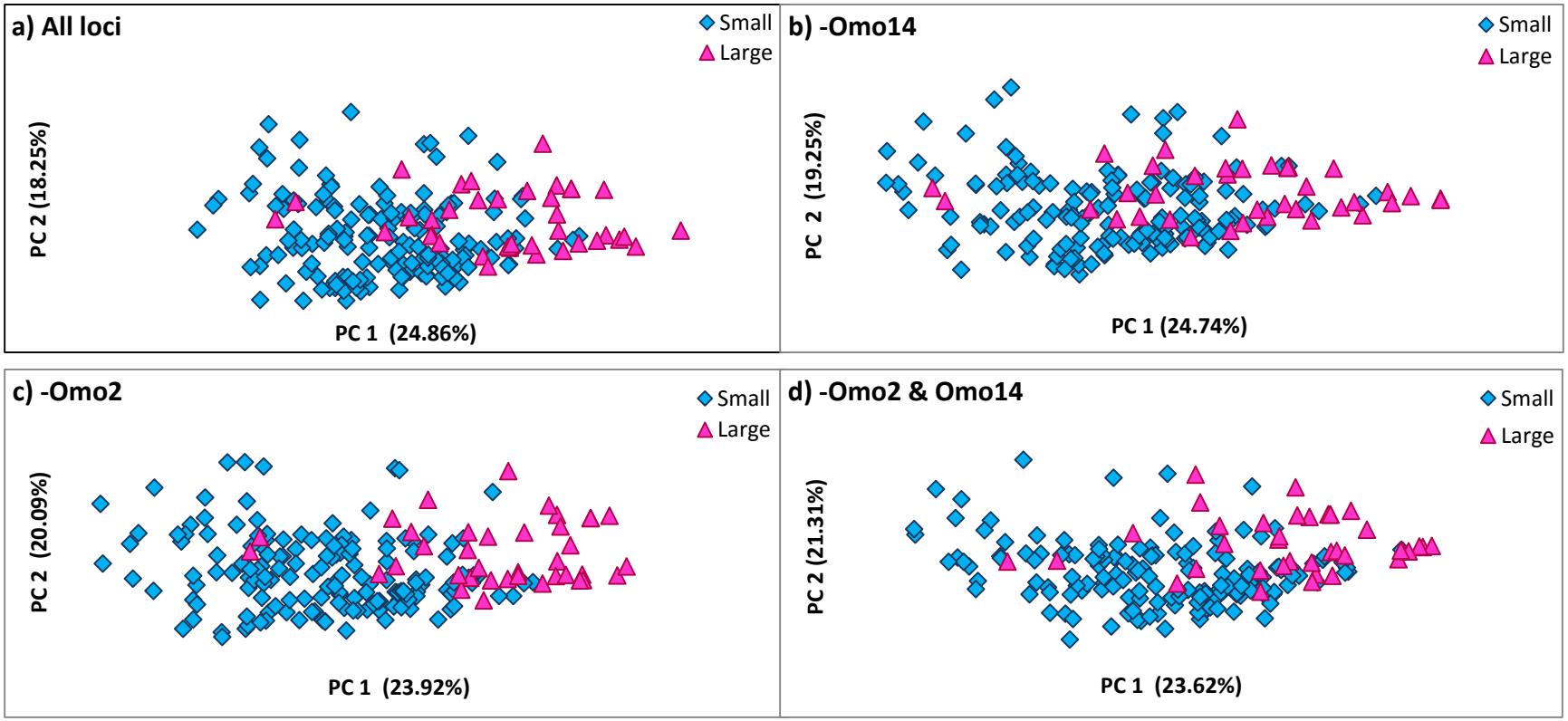


Figure 2.6. Principle coordinate analysis for adult smelt ecotypes collected in 2010/2011 (N=197, 159 small & 38 large morph). Results shown are a) all 7 microsatellite loci; b) 6 microsatellite loci with Omo14 removed; c) 6 microsatellite loci with Omo2 removed and d) 5 microsatellite loci with both Omo2 & Omo14 removed. All four plots show two dimensions of genetic variation; numbers in parentheses show percentage of genetic variation explained by one axis.

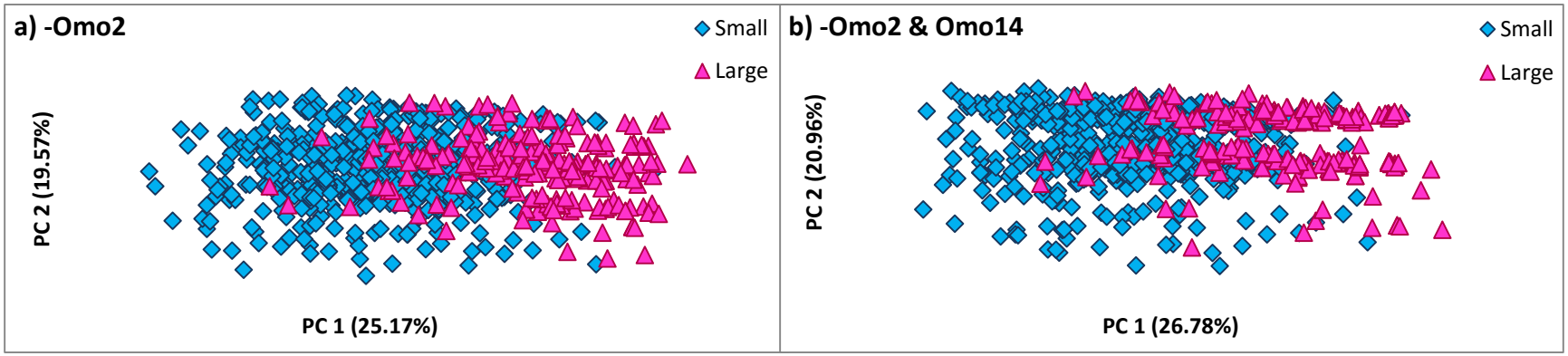
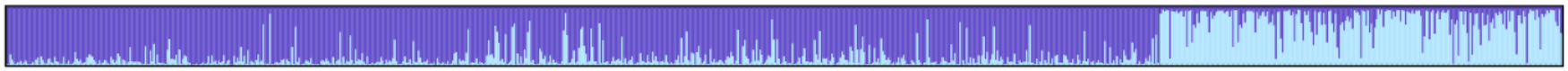


Figure 2.7. Principle coordinate analysis for adult smelt ecotypes collected in 2002, 2003, 2004, 2010 & 2011 combined (N=911, 675 small & 236 large). Results shown are a) 6 microsatellite loci with Omo2 removed; b) 5 microsatellite loci with Omo2 & Omo14 removed. Both plots show two dimensions of genetic variation; numbers in parentheses show percentage of genetic variation explained by one axis.



Small morph Large morph

Figure 2.8. Results of Bayesian analysis show population structure in the small and large morph of smelt. Each vertical line represents one individual and is partitioned into two coloured blocks that indicate the estimated membership of that individual into each of the two identified clusters.

Admixture coefficient, or q-values, were binned and plotted for both morphs. Individuals that had q-values between 0.0 and 0.1, 0.1-0.2, 0.2-0.3, etc. were binned together and the frequency of individuals belonging to each 'bin' was calculated. Although the two morphs were strongly associated with the two genetic clusters, there was overlap in their distributions of q-values (Figure 2.9a). Most (83.7%) small morph smelt had a q-value in the 0-0.2 range (65.5% < 0.1) and only 16.3% of small smelt had q-values greater than 0.2, to a maximum of 0.89 (1.3% >0.8). As for the large smelt, 63.1% had q-values in the 0.8-1.0 range (45.3% >0.9), and the remaining 36.9% had q-values less than 0.8, to a minimum q-value of 0.04 (4.2% <0.2). Similar patterns were seen with all loci combinations. In each case the majority of the small smelt had q-values in the 0.0-0.2 range, a smaller majority of the large smelt had q-values of 0.8-1.0, some individuals had intermediate q-values, and in all cases but one (small smelt with Omo2 removed) there were a few individuals with q-values strongly characteristic of the opposite morph (greater than 0.9 or less than 0.1) (figures 2.9b-d).

2.3.3 Gene flow between small and large smelt morphs

Estimated rates of gene flow showed some evidence of interbreeding between morphs. On average, for all loci combination and all years sampled, Bayesian estimates of gene flow obtained using BayesAss suggested that 0.5-4.4% of individuals categorized as small or large morph on the basis of their body size were actually derived from the opposing (genetically defined) morph (Table 2.2).

2.4. Discussion

Diversity in natural populations is often driven by adaptive radiations (Schluter, 2001; Bradbury *et al*., 2010), but the evolutionary forces driving differentiation can be difficult to

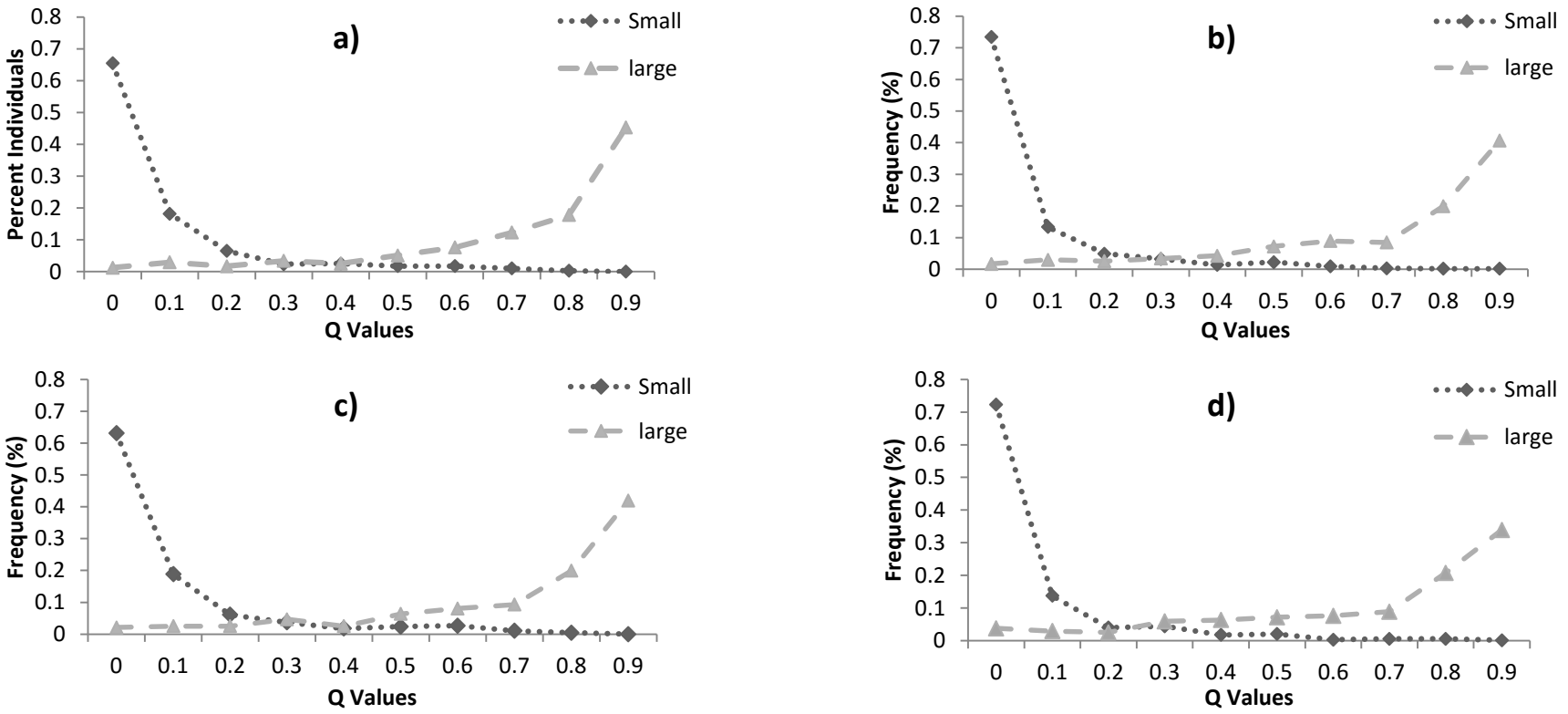


Figure 2.9. Frequency distribution of binned q-values from the STRUCTURE analysis for the two size classes (i.e., >185mm, and <155mm) where a) shows all 7 microsatellite loci b) shows 6 microsatellite loci with Omo14 removed, c) shows 6 microsatellite loci with Omo2 removed and d) shows 5 microsatellite loci with both Omo2 & Omo14 removed.

Table 2.2. Estimated rate of gene flow between the small and large morph, obtained using BayesAss. Values that differ significantly from zero, based on 95% confidence intervals, are in bold and underlined. Values marked with N/A are the same as the values below them, as Omo2 was not genotyped for 2002 and 2003 samples.

Loci Combination	Year→	2002		2003		2004		2010/2011		Combined	
	Gene flow from→ / Gene flow to↓	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
All Loci	Small	N/A	N/A	N/A	N/A	**0.988**	**0.005**	**0.98**	**0.028**	**0.995**	**0.009**
	Large	N/A	N/A	N/A	N/A	**0.012**	**0.995**	**0.02**	**0.972**	**0.005**	**0.991**
Omo2 Removed	Small	**0.991**	**0.029**	**0.96**	**0.014**	**0.983**	**0.005**	**0.967**	**0.03**	**0.993**	**0.01**
	Large	**0.009**	**0.971**	**0.04**	**0.986**	**0.017**	**0.995**	**0.033**	**0.97**	**0.007**	**0.99**
Omo14 Removed	Small	N/A	N/A	N/A	N/A	**0.984**	**0.005**	**0.976**	**0.032**	**0.995**	**0.008**
	Large	N/A	N/A	N/A	N/A	**0.016**	**0.995**	**0.027**	**0.968**	**0.005**	**0.992**
Omo2&14 Removed	Small	**0.991**	**0.024**	**0.956**	**0.016**	**0.981**	**0.005**	**0.957**	**0.036**	**0.994**	**0.009**
	Large	**0.009**	**0.972**	**0.044**	**0.984**	**0.019**	**0.995**	**0.043**	**0.964**	**0.006**	**0.991**

discern. There are several theories predicting that adaptive radiations in sympatric fish populations have evolved through disruptive selection for opposing niches (Carlson *et al*., 2009; Bradbury *et al*., 2010) followed by reproductive isolation, though depending on the degree of contact, the nature of both processes may differ. Divergence between individuals of the same species with regards to feeding specializations is well documented in the literature (Nosil, Vines & Funk, 2005; Barette, Daigle & Dodson, 2009; Bradbury *et al*., 2010). There are many documented cases of lacustrine fish species that have evolved into large, benthic specialists and small limnetic specialists, most of which suggest convergent adaptive evolution (Robinson & Wilson, 1994; Bradbury *et al*., 2010). Two studies (Taylor & Bentzen, 1993b; Bradbury *et al*., 2010) have previously documented phenotypic and genetic bimodality of the Rainbow Smelt in Lochaber Lake, observations that are consistent with postglacial ecological divergence (Bradbury *et al*., 2010). The small and large morphs of smelt in Lochaber Lake differed significantly in shape, size, and the number of gill rakers, differences that are associated to two different feeding specializations, findings that are consistent with morphological forms across the species range with regards to selection for similar niches (Bradbury *et al*., 2010). When compared to allopatric smelt found in Newfoundland using several phenotypic markers, the landlocked sympatric smelt in Lochaber Lake were found to have greater morphology differences than the allopatric smelt, suggesting larger character displacement and thus less overlap between the two feeding specialists (Bradbury *et al*., 2010).

The present study aimed to further evaluate the general pattern of genetic differentiation observed between the small and the large landlocked Rainbow Smelt in Lochaber Lake and its stability over time. All statistical analyses pointed to the presence of two genetically distinguishable populations of smelt, one comprised of phenotypically small smelt and the other

comprised of phenotypically large smelt. Moreover, the genetic differentiation between the small and the large morph of smelt in Lochaber Lake appears to be stable and continues to persist over the time period investigated here despite evidence of persistent bi-directional gene flow between the two morphs. Taken together these results suggest that there is evidence for disruptive selection and ecological speciation with regards to preferred trophic specialization.

Evidence of genetic divergence connected with adaptive specialization previously observed in Lochaber Lake smelt (Taylor & Bentzen, 1993a; Bradbury *et al*., 2010) is consistent with previous findings of isolation in sympatry. This study demonstrated that despite the spatiotemporal overlap in spawning for both small and large morph in the two streams, there is relatively strong genetic differentiation between the two morphs (F_{ST} = 0.070-0116). The degree of divergence observed in sympatry in Lochaber Lake is comparable to that of anadromous smelt populations that are separated in some cases by more than 2000km (Bradbury *et al*., 2006; Bradbury *et al*., 2010). Both Bayesian and PCoA analyses pointed to the presence of two distinguishable groups, although some overlap between the two morphs was observed. This overlap is likely due to gene flow, a conclusion supported by the fact that 0.5-4.4% of adult smelt genetically identified as the opposite morph despite their fork length at sampling time. Despite the overlap in spawning and evidence of gene flow, genetic divergence between the small and the large smelt persisted and was relatively stable over a 10 year period.

There are several studies comparing the genetic differences between recently diverged landlocked populations of Rainbow Smelt. Their F_{ST} ranges are typically smaller or comparable to those found in the Lochaber Lake smelt. For example, a sympatric pair of Rainbow Smelt in Lac Saint-Jean, Quebec that also have spatiotemporal overlap in spawning, had F_{ST}'s of 0.024 and 0.013, respectively, in two rivers tributary to Lack Saint-Jean (Saint-Laurent *et al*., 2003).

The Rainbow Smelt in that study were also found to all have derived from one glacial lineage, except for three single individuals (Saint-Laurent *et al*., 2003), much like the smelt in Lochaber Lake. Gene flow between the small and the large form in each river was found to be high, where the gene flow between morphs was found to be almost as pronounced as the gene flow within each morph between rivers (Saint-Laurent *et al*., 2003). Despite the similarities between the Lac Saint-Jean smelt and the Lochaber Lake smelt, the divergence between the two Lac Saint-Jean morphs was much less than what was observed in Lochaber Lake.

Results from a sympatric pair of smelt morphs in Lake Utopia, New Brunswick (Bradbury *et al*., 2011) revealed F_{ST}'s of 0 to 0.067, with an average F_{ST} of 0.026. In Lake Utopia however, the two morphs spawn in five different areas, the smaller morphs spawn in three different brooks and the larger morph spawning in two outlet streams that connect to other lakes (Bradbury *et al*., 2010), and therefore unlike the Lochaber Lake smelt, no spatiotemporal overlap exists in their spawning patterns. Bradbury *et al*., (2010) compared an unusual case of a small morph estuarine smelt population in Newfoundland, located in close proximity to several large morph populations, to those of Lochaber Lake. Fertilization experiments within and between the anadromous morphs showed that pure crosses had a 79% fertilization success, whereas hybrid crosses averaged 78% success, suggesting that gene flow is not inhibited between the small and large anadromous morphs. For the sympatric Lochaber Lake morphs in the same study, pure crosses had fertilization success similar to those in the anadromous crosses, but fertilization success in hybrid crosses declined by 30-50% (Bradbury *et al*., 2010). Genetic differentiation (F_{ST}) between the estuarine morphs was also compared to differentiation observed in small samples of the two Lochaber Lake morphs. The sympatric smelt from Lochaber Lake had an $F_{ST} = 0.076$, whereas the estuarine smelt had an $F_{ST} = 0.06$ (Bradbury *et al*., 2010).

Higher levels of differentiation are typically associated with sympatric pairs that stem from multiple invasions (Saint-Laurent *et al*., 2003). Saint-Laurent *et al*. (2003) posited that F_{ST} values smaller than 0.05 are associated with sympatric ecotypes that have been found to have intra-lacustrine divergence, whereas F_{ST} greater than 0.20 are associated with sympatric populations of whitefish, stickleback and trout that have evolved following multiple invasions.

Taken in combination with the results of earlier research (Taylor and Bentzen,1993b; Bradbury *et al*., 2010) it is evident that the genetic differentiation observed between the two smelt morphs in Lochaber Lake is persistent and stable. This in turn suggests a role for disruptive selection in the formation of the diversity observed in this postglacial smelt population.

There is also evidence of hybridization between the Lochaber Lake smelt morphs based on the STRUCTURE results. A previous lab cross experiment showed that the two smelt morphs are capable of cross-fertilization, although there was a decrease in fertilization success compared to fertilization within morph (Bradbury *et al*., 2010). This study found that individuals that are either phenotypically small or large were in some instances identified genetically as the opposite morph or as intermediate (>0.2, <0.8) based on their q-values. Intermediate q-values also support a hypothesis of continued gene flow between the two populations. Because age data is missing for this study, it is not known for instance, whether the morphologically small smelt that are genetically large might in fact spawn in a subsequent year as large smelt or whether they are hybrids.

Genetic differentiation linked with adaptive specialization observed in Lochaber Lake is consistent with recurrent isolation in sympatry. While the stability of morphological differences between sympatric morphs has been well documented, studies of genetic differentiation spanning

more than a single generation have rarely been done (Tessier & Bernatchez, 1999). By using samples that span more than one generation, one can infer that the genetic structure observed is stable and not simply an artefact of non-random sampling, stochastic temporal fluctuations in allele frequencies caused by genetic drift, or variations in reproductive success leading to changes in genotypic frequencies due to demographic stochasticity. Genetic data from samples of both morphs collected over the course of five non-consecutive years provide evidence that the significant genetic differentiation between the small and the large morph is stable and has continued to be stable across several generations. The same can be said for the sympatric smelt in Lake Utopia. Using data on 10 microsatellite loci, individual smelt consistently clustered into two distinct groups of phenotypically small and large morphs, a genetic divergence that persisted over a span of two decades despite the continued presence of gene flow and hybridization, which was visible with the persistent presence of genetic hybrids and phenotypically small smelt with a large morph genotype (Bradbury *et al*., 2010; Bradbury *et al*., 2011).

The hybridization rates observed in this study suggests limited but ongoing gene flow between the small and the large morph of smelt. The decline in fertilization success reported by Bradbury *et al*., (2010) supports the idea that reinforcement plays a significant role in sympatric isolation. The same conclusion has been made from prior studies in other species based on observations of increased prezygotic isolation in sympatry compared to that in allopatry (Coyne & Orr, 1997; Bradbury *et al*., 2010). Prior studies suggest that stronger prezygotic than postzygotic isolation is present in pairs of fishes that are recently separated (Coyne & Orr, 1989; Nosil, 2007; Funk & Nosil, 2008; Lowry *et al*., 2008, Bradbury *et al*., 2010) which supports the drop in fertilization success in Bradbury *et al*., (2010)'s reciprocal crosses

The rate of hybridization has been an important focus in other studies as well. Unlike the

Rainbow Smelt's presumed prezygotic isolation, the reproductive isolation in sympatric stickleback pairs seems to be driven by both prezygotic and postzygotic ecological processes (Hatfield & Schluter, 1999). However, stickleback hybrids appear to suffer lowered fitness through ecological mechanisms and less so through fertilization (Hatfield & Schluter, 1999).

In Lake Whitefish, cross fertilization between sympatric morphs was just as successful as fertilization within pure forms, but the hybrid embryos suffered a two- to four-fold higher mortality rates than those of pure crosses (Lu & Bernatchez, 1998). Significant divergence was observed between both the allopatric and sympatric pairs of smelt in Bradbury *et al*. (2010)'s study, and given that hybrid fertilization was successful, it is assumed that isolating processes other than fertilization barriers may play a role.

The persistence and the stability of sympatric ecotypes or incipient species of fish depends on the continuous balance of hybridization and the selective removal of hybrids (Bradbury *et al*., 2011). Several past studies have discussed the possibility that sympatric forms could merge ("reverse speciate") into a single population (Frost 1965; Coyne & Orr, 2004; Taylor *et al*., 2006; Bradbury *et al*., 2011; Baillie *et al*., 2016), a situation that may be most often associated with environmental change and young post glacial lineages (Bradbury *et al*., 2011). If hybrid fitness levels increase, or environmental cues associated with the maintenance of divergence change, the stability of a sympatric pair may be compromised (Bradbury *et al*., 2011). Based on data collected that span more than one generation, the stability and persistence of sympatric forms in both Lochaber Lake and Lake Utopia suggests a dynamic balance between hybridization and the selective removal of hybrid individuals.

The use of neutral genetic markers such as microsatellites to resolve ecological divergence has been questioned in the recent years (Gavrilets & Vose, 2005). There are many

studies that suggest that alleles at loci that are not linked to genes under direct selection can potentially flow freely right through the population, uninhibited by selection (Gavrilets & Vose, 2005). This microsatellite analysis shows that the movement of neutral alleles between the smelt morphs does occur, but is likely not enough for the collapse of the two populations into one. Because it is unlikely that the microsatellites used in this study are linked to genes under selection, the structure that is observed in this study likely reflects a strong barrier to gene flow.

There is however, potential for historical isolation to play a role in the Lochaber Lake pair, but this remains an uncertainty (Bradbury *et al*., 2010). Another hypothesis is that this isolation seen here in sympatry reflects historical vicariance and genetic drift followed by secondary contact, but there is currently no evidence supporting a longstanding historical subdivision between the ecotypes (Bradbury *et al*., 2010). The lack of evidence suggests that the divergence seen here is evolutionarily recent, likely having occurred during the present interglacial (Bradbury *et al*., 2010). Taylor and Bentzen (1993b) conducted mtDNA analyses and found that the two morphs in Lochaber Lake are more closely related to one another than either is to their corresponding morph in Lake Utopia, New Brunswick, implying that the two populations of smelt seen in Lochaber have diverged *in situ* from a common ancestral population that colonized the lake.

2.5 Conclusion

In summary, I examined the pattern of differentiation in recently derived morphs of Rainbow smelt in sympatry. Levels of morphological and genetic differences suggest that the two morphs of smelt are very despite spatiotemporal overlap in spawning and subsequent gene flow. This partial barrier to hybridization suggests that reinforcement through prezygotic incompatibilities plays a very important role in keeping the two morphs from dissolving into one large population.

However, because gene flow is possible, any environmental changes could impact the stability of these two populations in Lochaber Lake.

CHAPTER 3 Morph and hybrid identity of larvae in a landlocked sympatric pair of Rainbow Smelt (Osmerus mordax) in Lochaber Lake, Nova Scotia

3.1 Introduction

Understanding the process of species formation is one of the main goals in evolutionary biology. Because species formation is generally too slow to be directly observed, many speciation studies have depended on the comparison of closely related, but completely isolated species (Otte & Endler, 1989). However, there are drawbacks to such approaches, as they are more likely to provide information about the characteristics of these species rather than provide information about the processes that gave rise to them (McPhail, 1993). The study of populations still capable of gene exchange but showing the ability to remain as distinct gene pools could provide better insight into the early stages of species formation (Chouinard *et al*., 1996).

Many characteristics of sympatric fish ecotypes that are found in north temperate lakes make them of special interest for this purpose (Lu & Bernatchez, 1998). These sympatric forms are specialized for distinct ecological niches, and are often found in post-glacial lakes that have developed following the last glacial retreat, which places a maximum time of around 15 000 years for these differences to have developed (Lu & Bernatchez, 1998).

Lochaber Lake near Antigonish, Nova Scotia, harbours a sympatric pair of Rainbow Smelt (*Osmerus mordax*) ecotypes. The two ecotypes differ in phenotypic traits associated with the occupation of different trophic niches (Taylor & Bentzen, 1993b), and because of this it is generally believed that reproductive isolation between the two morphs is mainly driven by ecological divergent selection (Lu & Bernatchez, 1998). The small microphagous morph of smelt typically mature at 105-155mm fork length, have larger eyes, shorter upper jaws and more

gill rakers than the large macrophagous morph which matures at 185-300mm fork length (Taylor & Bentzen, 1993b). Both morphs spawn in the same two streams (Hurlbert and McNab Brooks), and even though a small difference in peak spawning date exists, both forms are found spawning in the same streams on the same nights (Taylor & Bentzen, 1993b, P. Bentzen pers. com.).

Experiments performed on Lochaber Lake smelts found that fertilization rates in pure crosses (i.e., within morphs) were approximately 80%, but declined to approximately 55-65% for hybrid crosses (Bradbury *et al*., 2010). Notwithstanding the lower fertilization rate, the results still demonstrated that fertilization between morphs can occur. Estimates of gene flow between the two morphs also revealed that 0.5-4.4% of individuals categorized as either small or large morph based on their body size were actually derived from the opposing (genetically defined) morph (see Chapter 2). Genetic variation (F_{ST}) between the small and large smelt in Lochaber Lake also suggest that the two morphs are highly differentiated (F_{ST} = 0.070-0.116) (see chapter 2). The observed genetic divergence between the sympatric small and large smelt is equivalent to the divergence expected for anadromous smelt populations separated by more than 2000km (Bradbury *et al*, 2006). The absence of spatial and complete temporal segregation during spawning, and apparent lack of mate choice mechanisms suggest that hybrids between the two morphs may be expected.

Bradbury *et al*., (2010) showed that despite some reduction in fertilization success, interbreeding between the two smelt ecotypes can occur. Taken together with the facts that the two forms overlap in time and location of spawning, and have external fertilization, we can thus assume that pre-zygotic isolation is not the reason behind the low (0.5-4.4%) number of hybrid adult individuals in Lochaber Lake, but instead post-zygotic isolation may be occurring. Post-zygotic isolation can be intrinsic or extrinsic, where the former is considered to reflect

developmental problems in hybrids regardless of the environment in which they are found, and arises from divergent developmental systems that do not cooperate within a single genome (Coyne, 1992; Rogers & Bernatchez, 2006). The latter is environmentally dependent, when hybrids inherit intermediate phenotypes that experience lowered fitness in specific environments, due to conflicting natural selection (Schluter, 2000). Such post-zygotic isolation can arise because intermediate phenotypes are less efficient at catching prey in the wild, or because they have intermediate defenses that leave them susceptible to parasitism and predation (Schluter, 2001). Hybrids may exploit their available resources (in this case, likely prey) much less effectively than the specialized small and large smelt (Schluter, 2001), thus preventing hybrids to reach adult stages. So long as no intermediate environment exists for the hybrids to thrive in, individuals with intermediate phenotype will be subject to divergent selection that acted on their parental forms (Schluter, 2000).

The scenario described above is a form of disruptive selection. In disruptive selection, extreme phenotypes, in this case small versus large sized smelt, gain a fitness advantage over intermediate phenotypes, thereby driving coexisting phenotypes further apart through natural selection until each type, or morph, resides on dissimilar fitness peaks (Ruefller *et al*., 2006). This may involve the fixation of different alleles that are beneficial in one environment but detrimental in the other (Schluter 2009). Hybrids may have intermediate phenotypes, leading to fitness that is reduced due to ecological selection pressures (Schluter, 2001).

Hybrid survivorship has been studied in sympatric pairs of sticklebacks (Hatfield & Schluter, 1996; Rundle & Schluter, 1998; Hatfield & Schluter, 1999; Rundle, 2002; Gow *et al*., 2007; Taylor *et al*., 2011). One study reported that laboratory raised hybrids between benthic and limnetic parents were behaviorally and morphologically intermediate relative to the parental

species in traits associated to foraging ecology (McPhail, 1984, 1992). Further, analyses of genetic cohorts have shown that the frequency of hybrid genotypes declines with age in natural stickleback populations, supporting the hypothesis that selection against hybrids plays a major role in the reproductive isolation observed between the stickleback pairs in the wild (Gow *et al*., 2007). In combination, these observations imply that there is an ecological component driving the genetic divergences and reproductive isolation observed between the benthic and limnetic sticklebacks (Taylor *et al*., 2011).

3.1.1 Chapter Objectives

Results from chapter 2 reveal that hybrids, though relatively scarce considering the shared spawning habitat of the two morphs, exist in the adult populations of both small and large morph of Rainbow Smelt. It is therefore possible that the patterns previously studied in the stickleback populations are also occurring in Lochaber Lake.

This chapter examines the genetic relationships between the adults of the small and large morph and larval smelt collected from Lochaber Lake over a period of two years. The objectives of this study were to (1) determine whether larvae could be identified to morph using microsatellite markers, (2) determine whether hybrid status (pure vs. hybrid) could be ascertained, and (3) if so, to test whether the frequency of hybrids differs between larvae and adult spawners.

3.2 Materials and Methods

3.2.1 Study Area

The study area consisted of the two smelt spawning streams, McNab and Hurlbert Brooks that flow into Lochaber Lake, Nova Scotia. See chapter 2 for adult smelt sampling and a map of Lochaber Lake.

3.2.2 Field Methods

Adult smelt were sampled during the spawning period in March and April in 2010 and 2011, in McNab and Hurlbert Brooks. All fish were caught after dark with the use of small dip nets. The ecotype and sex of each fish was noted and a small pectoral or caudal fin clip was taken from every smelt and preserved in tubes containing 100% ethanol until further use. A total of 216 small and 58 large morph smelt were sampled, genotyped and included in the dataset.

Larval smelt were sampled in 2010 and 2011. In both years, sampling efforts began in May and continued at monthly intervals until October. During the 2010 sampling season and the start of the 2011 sampling season, a small aluminum boat with a 10HP outboard motor was used to tow a 1 m diameter ring net with a mesh size of 250µm. Each tow was done in a circular fashion for 20 min at a constant speed of 3-4 km/h, as determined using a field GPS. From June 2011 and the subsequent months, a ½m by 1m neuston net with a mesh size of 650µm replaced the ring net which appeared to increase the catch rate of larvae. All net tows were conducted after dark. Everything caught in the net, including copepods was preserved in 500mL mason jars containing 100% ethanol. The larvae were later individually picked out and put into smaller vials containing 100% ethanol. A total of 1619 larvae were caught; however, out of the 1619 only 1123 were used for data collection and out of those only 348 were successfully genotyped because of problems with DNA quality (Table 3.1).

3.2.3 DNA isolation and data collection

Tissue samples from all 268 adults and 1123 larvae were digested using a mixture of digestion buffer and proteinase K. DNA was extracted using the glass milk protocol described by Elphinstone *et al.* (2003), and modified to work with a 96-well filter plate and automated using a Perkin Elmer MultiPROBE II liquid handling system. Extracted DNA was stored at -

20ºC in preparation for further analysis. Ten neutral polymorphic microsatellite loci developed for *O. mordax* were used as follows: *Omo1, Omo2, Omo3, Omo4, Omo5, Omo9, Omo11, Omo13, Omo14 & Omo15* (Coulson *et al.*, 2006).

Individuals were genotyped using polymerase chain reaction (PCR) in 5- or 10- µL volumes containing 20-100ng DNA, 1.5mM $MgCl_2$, 80µM of each dNTP, 0.5 U *Taq* DNA polymerase (Applied Biosystems), 0.3 µM of each primer if the forward primers were end labeled with HEX dye or 0.1µM of forward primer, 1µM of reverse and 1µM of a HEX labeled tailed primer and 1x PCR buffer (10mM Tris-HCl, pH8.3; 50mM KCl). Two temperature profiles were used for touchdown PCRs.

Touchdown PCR was used for all primer pairs; OmoTD2 program was used for Omo1, Omo2, Omo3, Omo4 and Omo5 and OmoTD5 was used for Omo9, Omo11, Omo13, Omo14 and Omo15. Touchdown PCR conditions were as follows for OmoTD2: 94ºC for 2 min, followed by five cycles of 94ºC for 30 s, 66ºC for 30 s (-1ºC for the following four cycles), 72ºC for 30 s, followed by 25 cycles of 94ºC for 30 s, 62ºC for 30 s and 72ºC for 30 s with a final extension was held at 72ºC for 5 min. For OmoTD5, PCR conditions were as follows: 94ºC for 2 min, followed by four cycles of 94ºC for 30 s, 64ºC for 30 s (-1ºC for the following three cycles), 72ºC for 30 s, followed by 26 cycles of 94ºC for 30 s, 60ºC for 30 s and 72ºC for 30 s with a final extension was held at 72ºC for 5 min. Reactions were run on Eppendorf thermocyclers and visualized using a denaturing polyacrylamide gel electrophoresis on an FMBioII system (Hitachi Genetic Systems). Alleles were scored by hand in reference to a molecular weight size standard ladder incorporated into each gel. Every image also incorporated a positive control and a redundant sample in order to allow comparison of new samples to those of known allele sizes

Table 3.1. Date and number or larval smelt caught in 2010 and 2011 versus the number of larval smelt genotyped.

Date	Area Sampled	Life history stage	Caught	Genotyped
May 11 2010	Lochaber Lake	larvae	70	44
May 12 2010	Lochaber Lake	larvae	23	18
May 17 2010	Lochaber Lake	larvae	37	30
May 18 2010	Lochaber Lake	larvae	191	186
May 27 2010	Lochaber Lake	larvae	9	8
June 30 2010	Lochaber Lake	larvae	4	1
August 12 2010	Lochaber Lake	larvae	9	5
September 16 2010	Lochaber Lake	n/a	0	0
October 17 2010	Lochaber Lake	n/a	0	0
May 20 2011	Lochaber Lake	larvae	402	56
June 29 2011	Lochaber Lake	larvae	442	0
July 20 2011	Lochaber Lake	larvae	432	0
Total			**1619**	**348**

which helped ensure consistent scoring across both ecotypes and loci.

3.2.4 Data Analyses

Microsatellite scores were checked to ensure that common genotyping errors, such as null alleles, stutter and large allele drop out were not present using MICRO-CHECKER v2.2.3 (van Oosterhout *et al*., 2004). For each locus and both ecotypes the total number of alleles, the size range of the allele, expected- (H_E) and observed (H_O) heterozygosity and inbreeding coefficient (F_{IS}) were calculated using GenAlEx v. 6.41 (Peakall & Smouse, 2006); allelic richness (A_E) for all loci for both adult morphs and the larval smelt, as well as the average microsatellite heterozygosity for each morph were found using FSTAT v. 1.2 (Goudet, 1995). Deviations from HWE were assessed using the Markov Chain Monte Carlo (MCMC) approximation of Fisher's exact test and linkage disequilibrium (LD) was assessed for all possible loci combinations using the MCMC simulated exact tests as implemented in ARLEQUIN 3.5 (Excoffier & Lischer, 2010). In tests for both HWE and LE, statistical significance (α) was adjusted for the number of simultaneous tests k (α/k for $\alpha = 0.05$) using a sequential Bonferroni procedure (Rice, 1989) in order to reduce Type I errors. Because the action of selection can generate patterns of LD and violates some of the critical assumptions of HWE, both ecotypes were evaluated separately. Loci were removed only if LD or Hardy-Weinberg disequilibrium was detected for both ecotypes. These results are reported in Appendix 3.1.

Estimates of pair-wise F_{ST} between the two ecotypes and the larvae, and their statistical significance, were calculated using Microsatellite Analyser (MSA) (Dieringer & Schlötterer, 2003). Principal Coordinate Analyses (PCoA) were carried out on all individuals using a distance matrix created from the codominant genotypic distance algorithm of GenAlEx v.6.41 (Peakall & Smouse, 2006, algorithm described in Smouse & Peakall, 1999).

Bayesian clustering analysis was conducted in STRUCTURE v.2.3.3 (Pritchard *et al*., 2003) and tested the presence of multiple discrete populations. This method uses assumptions of linkage disequilibrium and HWE among loci, establishes population structure, and uses a MCMC algorithm to assume populations that are not in linkage disequilibrium and develop estimates of the number of populations (K). The program was run four times for each value of K to ensure convergence of values, and with a burn-in of 100,000 reps, 500,000 reps after burn-in, and K ranging from 1 to 5. The most likely K value was found based on the rate of change of the log probability of data (ΔK) (Evanno *et al*., 2005) using STRUCTURE HARVESTER (Earl & von Holdt, 2011). STRUCTURE also estimates q-values, or admixture coefficients, for each individual, which represent the portion of each multilocus genotype allocated to a certain cluster or population. If hybrids between the two ecotypes occur, they are expected to possess an intermediate q-value. Results from replicate STRUCTURE runs were merged using CLUMPP 1.1.2 (Jakobsson & Rosenberg 2007) and graphically summarized using DISTRUCT 1.1 (Rosenberg 2004).

The rate of gene flow between the two morphs was estimated using a Bayesian approach implemented in BAYESASS v.3.0.1 (Wilson & Rannala, 2003). This was done for each year separately, and for the entire data set combined. BAYESASS was run with a burn-in of 1 million MCMC iterations followed by 10 million sample iterations. This process was done several times for each data set, each time with a different starting seed number to ensure consistent and accurate results, as recommended in the program manual. The results from all the runs were examined to make sure that there was convergence, but only the results from the first BAYESASS run were reported. Standard deviations given by the program were used to construct 95% confidence intervals around each estimated rate of migration. If the interval did

not include zero, migration rates were considered significant.

3.2.5 Individual assignment tests

The ability of each dataset to assign individuals to the most likely ecotype of origin was assessed using the method of Rannala and Mountain (1997) implemented in the genetic stock (GSI) program, ONCOR (Kalinowski *et al*., 2007). "Pure" individuals of each form were selected based on their q-values given by STRUCTURE. If an individual had a q-value smaller than 0.2 or greater than 0.8, it was considered a pure form of the ecotype associated with that particular end of the q value range. First, a baseline file consisting of 36 'pure' small morphs and 36 'pure' large morphs was created. Only 36 individuals were chosen due to the small sample size of large smelt. This file contains genotypes from populations that have potentially contributed to individuals in a 'fishery'. Then, a mixture file containing the genotypes of unknown larvae was compared to the baseline data, and using maximum likelihood, the mixture proportions of the larvae were estimated. Using the same baseline and mixture file, individual assignment for each larva was also estimated, using both genotype frequencies and mixture proportions.

3.2.6 Bayesian tests for hybridization

Two different Bayesian approaches were used to deduce the presence, as well as the frequency of several hybrid classes in the sampled smelt. As previously mentioned, Bayesian clustering was performed using STRUCTURE to estimate the admixture coefficients (q-values) for every individual. In addition to STRUCTURE, NEWHYBRIDS (Anderson & Thompson 2002) was used to examine hybridization and the potential occurrence of specific hybrid classes. NEWHYBRIDS uses a Gibbs sampler and Markov Chain Monte Carlo to estimate the posterior probability that genetically sampled individuals fall into a certain hybrid category ranging from

pure groups to F1, F2, and both backcrosses. NEWHYBRIDS was run with a burn-in of 100,000 iterations followed by 250,000 iterations after burn-in. Individuals that classified to a certain group with a probability >50% were considered correctly assigned.

In order to evaluate the capacity of both STRUCTURE and NEWHYBRIDS to correctly identify the several hybrid classes, HYBRIDLAB 1.0 was used to simulate individuals of each class (Nielsen *et al*., 2006). All six groups of pure and hybrid individuals (pure small, pure large, F1 hybrids, F2 hybrids and their reciprocal backcrosses) were simulated using the same 36 pure individuals of each group used in the ONCOR analyses. Using these individuals, HYBRIDLAB was used to simulate 100 pure individuals of each form and 100 individuals for each hybrid class. These individuals were then used for STRUCTURE and NEWHYBRIDS analyses using the same parameters and methods for each described above. The results were plotted in order to visualize how well the two programs identified actual simulated hybrids.

3.3 Results

3.3.1 Data quality

MICROCHECKER found evidence that Omo13 had null alleles in small, large and larval smelt, and therefore it was not used in data analyses. Arlequin results with subsequent Bonferroni corrections revealed no evidence of linkage between any loci. Results from Arlequin also revealed that Omo11 was not in HWE in the large smelt morph. As for the larval smelt, Omo4, Omo11 and Omo15 were not in HWE, which is likely due to a Wahlund effect as both morphs of smelt are present in the larval samples.

3.3.2 Genetic Diversity

The average microsatellite heterozygosity of the large smelt (0.52) was significantly different from that of the small smelt (0.74) (ANOVA, p=0.014) and that of the larval smelt

(0.72) (ANOVA, p=0.0.27) but the small smelt and the larvae means were not significantly different (ANOVA, p=0.76). The average number of alleles followed the same pattern; average number of alleles for large smelt (Na = 6.44) was significantly different from that for small smelt (Na = 11.56) (ANOVA, p=0.014) and that for larvae (Na = 12.44) (ANOVA, p=0.002), but the average number of alleles for larval smelt and small smelt were not significantly different (ANOVA, p=0.681).

Pairwise F_{ST} values compared between the small morph, large morph and the larval smelt were all statistically significant indicating genetic differentiation between the two morphs and the larvae. Pairwise F_{ST} was 0.067 ($p < 0.01$) when the small and the large morph were compared, 0.012 ($p < 0.01$) when the small morph and larvae were compared and 0.061 ($p < 0.01$) when the large morph and larvae were compared.

Principle Coordinate Analysis showed differences between the small and the large smelt, but also overlaps between the genotype distributions for the two morphs (Figure 3.1). The small morph appeared to be more genetically variable than the large morph and overlapped the more constricted genotypic distribution of the large morph. The distribution of genotypes for the larval smelt overlapped almost the entire genotype distributions for small and large smelt, but there were also many larvae with genotypes that fell outside of the space delimited by the genotypes of either adult morphs (Figure 3.1).

Bayesian clustering implemented in STRUCTURE supported the presence of two discrete groups of smelt, with K = 2 being the most likely number of clusters determined by the tests implemented in STRUCTURE HARVESTER (Appendix 2) (Figure 3.2). One cluster was predominant in adult small morph smelt and the other cluster was predominant in large morph smelt. Both genetic clusters were evident in the larval smelt.

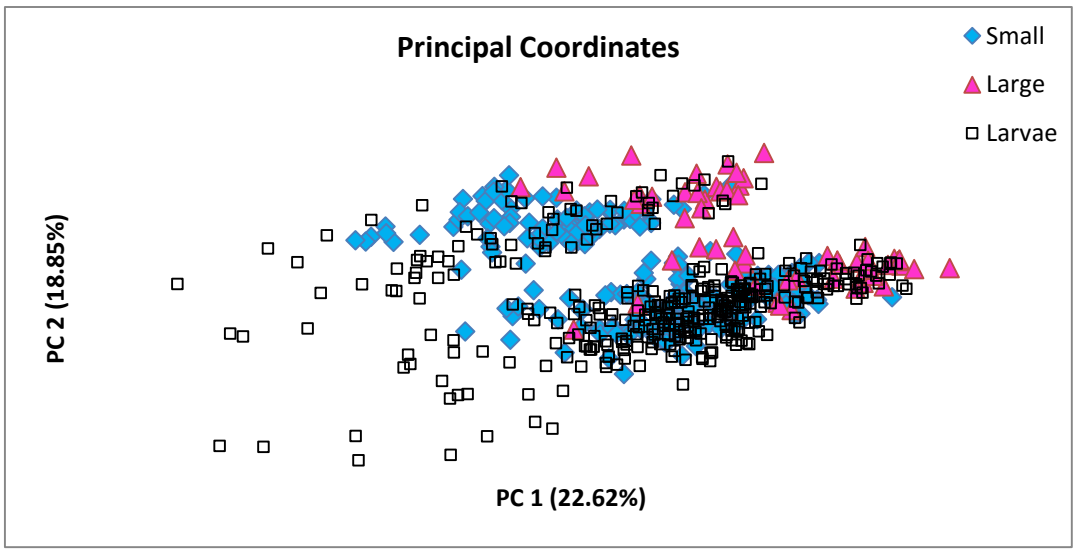


Figure 3.1. Principal Coordinate Analysis of genetic distance between the 216 small morph, the 58 large morph and the 348 larvae smelt collected in 2010 and 2011 and genotyped at 9 microsatellite loci. Two dimensions of genetic variation are shown, describing 41.47% of the total variation. The two clusters observed are due to having removed Omo13 because of null alleles.

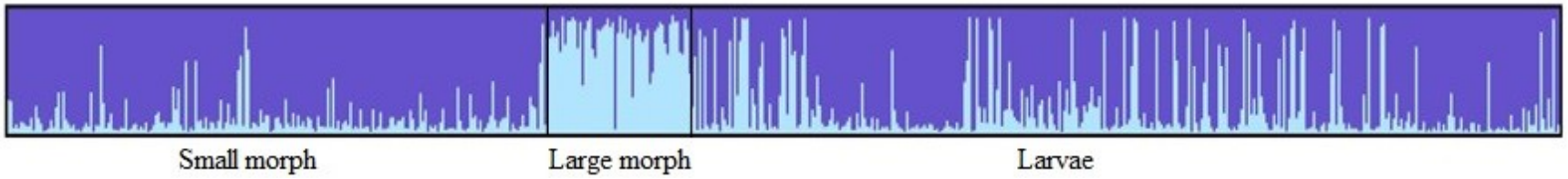


Figure 3.2. Bayesian clustering analysis of small and large morph as well as the larvae collected in 2010 and 2011. Each individual is represented by a single vertical line, partitioned into two coloured segments to indicate the estimated membership of that individual into each identified cluster.

Admixture coefficient, or q-values, were binned and plotted for both morphs and for larvae. Individuals that had q-values between 0.0 and 0.1, 0.1-0.2, 0.2-0.3, etc. were binned together and the frequency of individuals belonging to each 'bin' was calculated. The q-value distributions differed among the two morphs and the larvae, although the distributions tended to overlap in the intermediate q-value range (Figure 3.3). Most (83.8%) small morph had q-values in the 0-0.2 range (63.4% < 0.1). The remaining 16.2% of small smelt had q-values ranging up to a maximum of 0.88. Most (62.1%) large smelt had q-values in the 0.8-1.0 range (20.7% >0.9); the remaining 37.9% had q-values ranging down to 0.034. The q-value distribution for larval smelt was intermediate compared to those of the two size morphs, but was most similar to that of the small morph; 69.8% had q-values between 0.0 and 0.2 (54.9% <0.1), 30.2% had q-values greater than 0.2, including 10.1% that had q values in the 0.8-1.0 range.

A chi-square test based on binned q-values showed that the q distributions of larval and small morph smelt were significantly different (X^2=16.919, df=9, $p < 0.001$). A second q-value plot provided a clearer view of the difference between the larval q-values and the small smelt q-values (figure 3.4). In this case the q-values were arranged from smallest to largest for each ecotype then standardized to a common sample size to better show the trend regardless of the number of individuals for each class. The q-value distributions of small and large smelt were very different. The q-value distribution for the larval smelt fell between those of the adult groups, but was relatively more similar to that of the small morph; however, a marked difference in the q-value distributions of larval and small morph smelt at q-values greater than 0.6 provided evidence of the presence of large morph smelt in the larval collections.

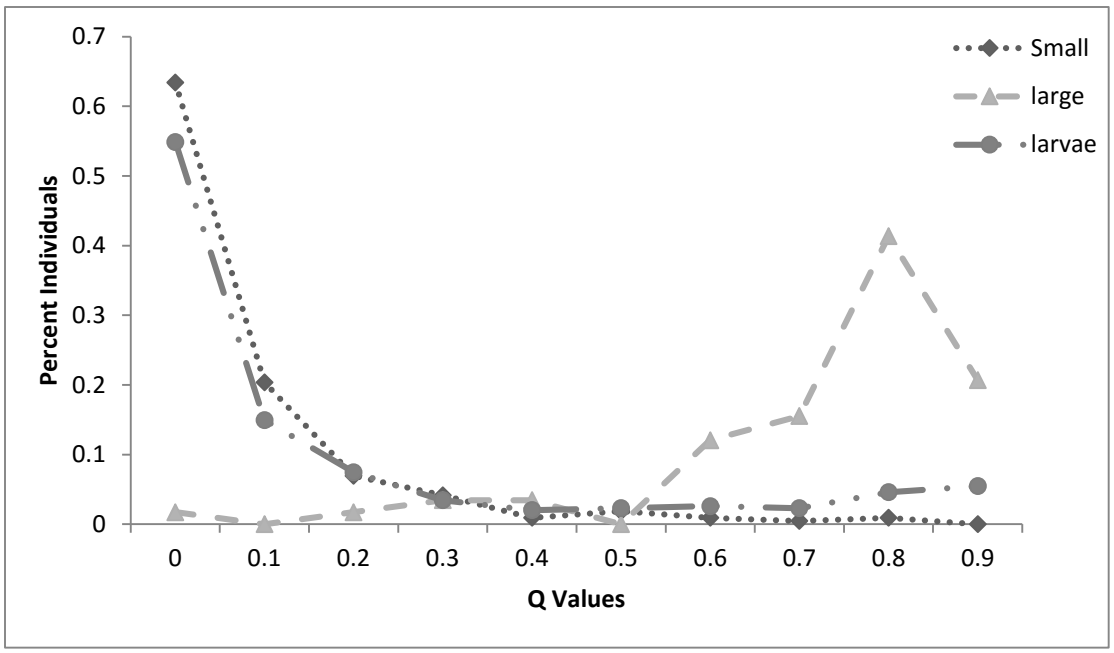


Figure 3.3. Frequency distribution of binned q-values obtained from the STRUCTURE analysis for the two adult morphs and the larval smelt.

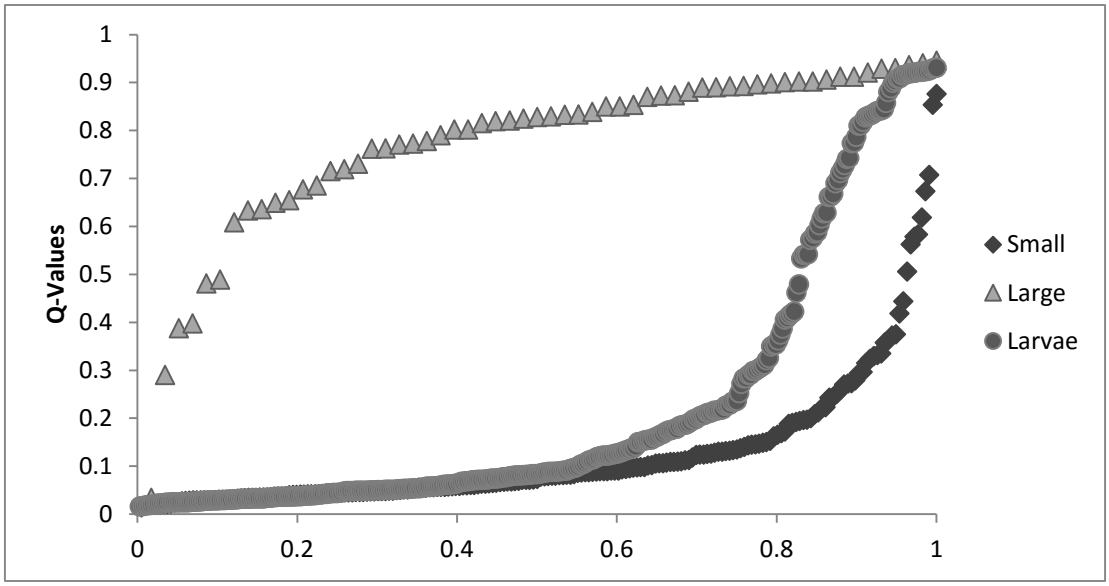


Figure 3.4. Q-values of for all individuals of small and large morphs as well as larval smelt sorted from smallest to largest. The distributions were scaled to a common sample size in order to better compare the shapes of the q-value distributions among groups of smelt.

3.3.3 Gene flow between small and large smelt morphs

Estimated rates of recent gene flow obtained using BayesAss showed evidence of low levels of gene flow from both the small smelt population into the large smelt population and vice versa. The estimated rate of gene flow from the small morph population into the large morph population was 2.4% (SD = 0.014, CI = 0.004-0.056); whereas, gene flow from the large morph population into the small morph population was estimated to be 1.9% (SD = 0.013, CI =0.003-0.052).

3.3.4 Individual Assignment Test

A baseline file containing 36 'pure' small and 36 'pure' large adult smelt was made based on q-values given by STRUCTURE. Small morph smelt with $q < 0.2$ and large morph smelt with $q > 0.8$ were defined as 'pure' representatives of their morph. Using these baseline data ONCOR's mixture analysis determined that approximately 82% (95% CI, 0.727, 0.863) and 18% (95% CI, 0.137, 0.273) of the larvae belonged to the small and large morph populations, respectively. Comparing these results to the q-values obtained from STRUCTURE (Figures 3.3 and 3.4), we can see that a large proportion of larvae do in fact have genotypes similar to those found in the small morph. Results from the individual assignments estimated by ONCOR showed that 283 (81.3%) of the 348 larvae were small morph, 54 (15.5%) were large morph and 11 (3.2%) did not clearly identify with either of the two morphs (Appendix 3). In comparison, using the same baseline data as above, ONCOR's mixture analyses using the baseline of 'pure' adults determined that approximately 96.14% (95%CI, 0.858, 0.994) and 3.86% (95% CI, 0.006, 0.142) of small smelt adults belonged to the small morph and the large morph respectively and that 97.71% (95% CI, 0.874, 1.000) and 2.29% (95% CI, 0.000, 0.126) of the large smelt adults belonged to the large morph and small morph respectively. Results from the individual

assignment test estimated by ONCOR showed that 19 (8.8%) of the 216 small morph identified as large morphs and 6 (10.3%) of the 58 large morphs identified as the small morph.

3.3.5 Hybridization analysis

The power to detect hybrid individuals and specific hybrid classes was evaluated using simulated hybrid individuals (see methods). Bayesian clustering using STRUCTURE (Figure 3.5) revealed that the simulated 'pure' morphs could be easily identified and distinguished from F1 and F2 hybrids, but that F1 and F2 individuals could not be differentiated from one another as both of these are characterized by intermediate admixture coefficient, not making it evident which type of hybrid they are characterized as. As for the backcrossed individuals, they displayed on average higher and lower admixture values, and also showed overlap with other classes, making it less obvious that they are hybrids. Using a relaxed threshold q-value of 0.5, Bayesian clustering using NEWHYBRID correctly identified all of the simulated pure small morphs and all of the simulated pure large morphs (Figure 3.6). Had a more stringent threshold of <0.2 and >0.8 been used, six of the simulated pure small morphs (16.6%) and seven of the simulated pure large morph (19.4%) would be considered wrongly identified, however when comparing these results to the graphical representation of 'pure' small and large morphs as seen in Figure 3.6 all pure forms are in fact correctly identified. Using a threshold q-value greater than 0.5, 65% of the simulated F1 hybrids were correctly identified, with the remaining 35% incorrectly identified as either the small or the large morph. NEWHYBRID lacked the power to detect F2, BC1 and BC2 hybrids, as no individuals simulated for those hybrid classes were correctly identified. Given the success rate observed for the identification of the pure forms and at least one of the hybrid classes, NEWHYBRID analysis was conducted on the 2010 and 2011 individuals. In this case, because the power to detect hybrids beyond the F2 class was lacking, q-

values obtained from all hybrid classes were summed and the final value was used to identify hybrids from pure forms. Using the same 0.5 q-value threshold as above, of 216 small morph adults, six (2.7%) were identified by NEWHYBRID as large morph, and seven (3.2%) were identified as hybrids (Figure 3.7). For 58 large morph adults, three (5.2%) were classified by NEWHYBRID as small morph, and four (6.7%) was identified as a hybrid. Out of the 348 larvae, 285 (81.9%) were identified as small morph, 53 (15.2%) as large morph, and ten (2.9%) as hybrids. See table 3.2 for combined results.

3.4 Discussion

The main objective of this study was to test whether smelt larvae could be identified to morph by using microsatellite markers, and whether or not hybrids could be identified among the larval smelt and if they are present, if they have a higher frequency than the hybrids observed in the adult spawners. Based on the results obtained in this study, unknown (larval) forms of smelt can be identified to morph when using microsatellite markers. The proportions of larval smelt genetically assigned to small or large morph, 82% and 18%, respectively as determined by ONCOR, is comparable to the phenotypic ratio observed in the adult populations (pers. obs.), where the small morph is present in much higher numbers than the large morph on any spawning night. Although the power to detect hybrids beyond F1's was weak, total q-value results summed for all hybrid classes still show that hybrids make up, between 3.2% and 6.7% of the adult small and large morph populations, as well as approximately 3% the larval samples, indicating that barriers to successful hybridization operate.

Hybrid studies done on sticklebacks in Paxton Lake, British Columbia, showed increased hybrid viability and fertility in a laboratory setting, but when introduced to pond and lake enclosures a significant reduction in the proportion of hybrid survival was found (Vamosi *et al*.,

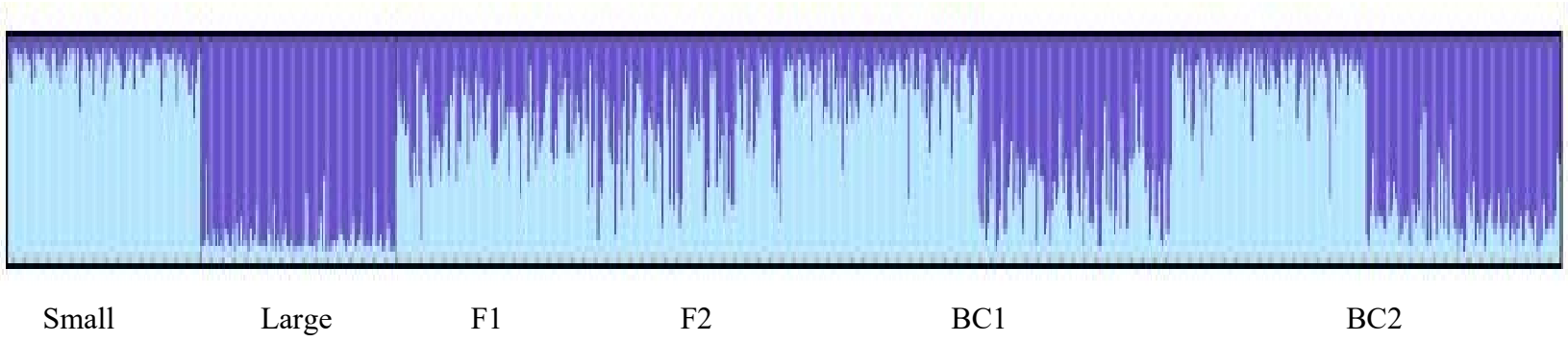


Figure 3.5. Bayesian assignment of simulated pure and F1, F2, BC1 and BC2 hybrid classes of smelt using STRUCTURE with K=2. Individuals were simulated using HYBRIDLAB.

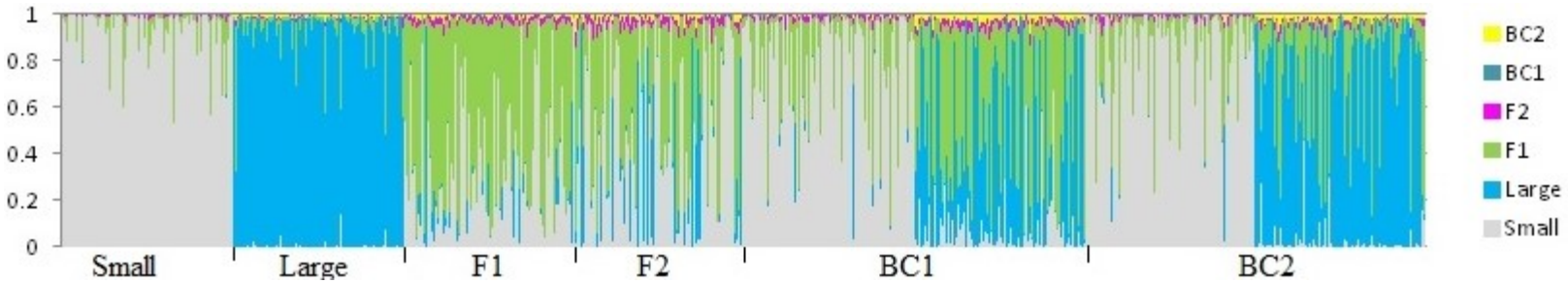


Figure 3.6. Bayesian assignment of simulated hybrid classes using New Hybrids with six possible hybrid classes of individuals (pure small, pure large, F1, F2, BC1 and BC2). Individuals were simulated using HYBRIDLAB.

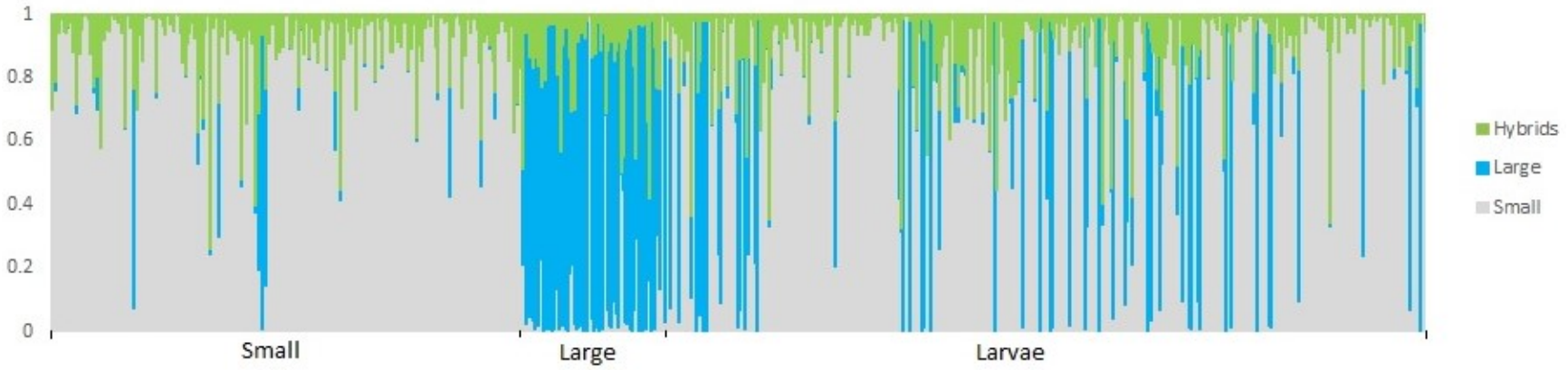


Figure 3.7. Bayesian identification of hybrid classes of the two adult morphs and the larval smelt collected in 2010 and 2011 using New Hybrids.

2000). It was estimated that 11.1% of stickleback hybrids in pond and lake enclosures were selected against per year (Vamosi *et al*., 2000). A similar study conducted on hybrids between two cyprinid species, the common shiner *Luxilus cornutus* (Mitchill) and the striped shiner *L. chryosocephalus* (Rafinesque) showed that 9.2% of hybrids were selected against per year (Dowling & Moore, 1985), showing evidence for strong selection against hybrids. The persistence and stability of the Lochaber Lake smelt pair is likely dependent on the selective landscape as changing biotic and abiotic conditions could unbalance the complex, which has been observed in a stickleback pair (Taylor *et al*., 2006). Further investigation into smelt hybrid viability in a laboratory setting and in the wild will be required in order to better understand the survival rate and the rate of selection currently occurring in the Lochaber Lake population.

Studies conducted in the wild showed that genetic intermediates between benthic and limnetic stickleback morphs from Paxton and Priest Lake in British Columbia steadily declined across successive life-history stages (Gow *et al*., 2007). In Priest Lake, hybrids suffered an 80% decline, from 19.6% of juvenile individuals being hybrids, to 3.7% of adults being hybrids. Paxton Lake hybrids suffered a 30% decline, starting out at 6.58% of juveniles being hybrids and declining to 4.7% hybrids in the adult population (Gow *et al*., 2007). Juvenile hybrid frequencies fluctuated considerably, whereas those at the adult level were relatively constant (Gow *et al*., 2007). These results support the prediction that selection against hybrids plays an important role in reproductive isolation of these ecotypes (Gow *et al*., 2007).

A study of hybrid fitness involving two allopatric whitefish populations belonging to two distinct glacial races was also done. Lu & Bernatchez (1998) conducted fertilization experiments between small morph Acadian and large morph Atlantic-Mississippian Lake Whitefish found in two different lakes. Crosses consisted of pure small crosses (small female x small male), pure

Table 3.2. Results obtained from most of the data analyses showing the same trend between the small, large and larval smelt.

	STRUCTURE		ONCOR		NEWHYBRID
PHENOTYPE	q <= 0.2	q >= 0.8	Ind. Ass.	mixture	
SMALL	83.80%	0.93%	91.2% small	96.14% small	94% small
			8.8% large	3.86% large	2.8% large
					3.2% hybrid
LARGE	1.72%	62.10%	10.% small	2.29% small	5.2% small
			89.7% large	97.71% large	87.9% large
					6.9% hybrid
LARVAE	69.80%	10.10%	81.3% small	82% small	81.9% small
			15.5% large	18% large	15.2% large
			3.2% unknown		2.9% hybrid

large crosses (large female x large male), and two hybrid crosses (small females x large males and large males x small females). High fertilization success was observed for all four crosses with success rates of 94.2%, 96.3%, 93.1% and 90.7% respectively. In contrast to this however, a higher embryonic mortality rate was observed in hybrid crosses compared to pure group crosses, where hybrids mortality rates were 2.7-4.7 times greater than those of pure crosses. Embryonic mortality was found to vary with each developmental stages, but highest mortality rates were observed between days 28 and days 44 after fertilization for both pure and hybrid crosses, however hybrid survival rate after 100 days post fertilization were still 45-60% for small female x large male crosses and large female x small male crosses respectively. Increased hybrid embryonic mortality supports a partial post-mating reproductive isolation hypothesis (Lu & Bernatchez, 1998); however, because this experiment showed a lack of evidence for reduced intrinsic viability and fertility of inter-ecotype hybrids observed in a laboratory setting, we can conclude that genome incompatibility is not the reason for post-mating isolation between the ecotypes (Schluter, 1996). Instead, there is evidence that supports the hypothesis of reproductive isolation is driven by divergent selection causing hybrid inferiority (Schluter, 1996).

Hybridization levels were also examined between the Arctic char (*Salvelinus alpinus*) and Dolly Varden Char (*Salvelinus malma*) in the Wood River Lake system in southwestern Alaska found that the pair are genetically distinct from one another. However, it was discovered that F1 hybrids were relatively rare, but the presence of post-F1 individuals indicated that those hybrids that are produced are viable and fertile (May-McNally *et al*., 2015). It is hypothesised that ecological factors may restrain the hybridization opportunities between these two particular species (May-McNally *et al*., 2015). For example, Dolly Varden and Arctic Char appear to use distinct spawning habitats in this lake system; Arctic Char use the mouths of creeks where there

are large submerged gravel beaches, whereas Dolly Varden have only been observed in two particular streams (Wood River and Yaho Creek) (May McNally *et al*., 2015). It has also been observed that their peak spawning times differ; Dolly Varden spawn around the third week of September and Arctic Char spawn in mid- to late October (May-McNally *et al*., 2015). They also noted that the Dolly Varden are rare in the lakes, which would reduce the levels of hybridization, much like the smelt morphs observed in Lochaber Lake where the small morph is much more numerous on any given day than the large morph. The level of hybridization is to some extent dependent on the relative density of the parental species (Wirtz, 1999; May-McNally *et al*., 2015). Based on the number of adult small versus the number of adult large smelt caught on any given year, this could potentially highly affect the number of hybrid larvae observed in the present study. As May-McNally (2015) claims, the lower abundance of one morph (Dolly Varden Char) likely limits interspecific encounters during the reproductive period and may influence to some extent the hybridization rates.

There also exists the possibility that hybrid smelt larvae are more abundant than detected in this study. Of the 1619 smelt larvae that were sampled, genotyping was attempted on 1123, but only 348 were successfully genotyped due to problems with DNA quality. It is possible that better genotyping success might have led to identification of a higher proportion of hybrids.

Although information about hybrid viability and survival rate is lacking in order to confidently say which processes are occurring in the smelt populations of Lochaber Lake, there are similarities between the sticklebacks and the smelt. Adult sticklebacks share the littoral zone during breeding season (Bentzen *et al*., 1984) much as the small and large morphs of smelt spawn in the same two streams at overlapping times in Lochaber Lake. The proportion of hybrids in adults also appears to be stable in the smelt, much like the adult sticklebacks.

Although the proportion of hybrids found in the larval smelt was much smaller than that found in juvenile sticklebacks, the sample sizes were also much smaller in the smelt than those used in sticklebacks, which could potentially result in lowered hybrid frequencies (see Table 3.1). When analyzing the power to detect smelt hybrids, there was a 35% decline in the power to detect F1's when compared to the detection of pure small and large adult morphs, suggesting that perhaps more hybrids are present in the larval smelt but the power to detect them all was lacking. A study that compared the power of microsatellites and single nucleotide polymorphisms (SNPs) for detecting hybrids in birds found that using 20 microsatellite loci outperformed nine SNPs in the identification of species, but were poorer in hybrid detection, suggesting that the most efficient and accurate way to identify individuals was to use a combination of both microsatellite loci and SNPs (Väli *et al*., 2010).

Despite evidence that several pairs of sympatric fish produce hybrids that are subsequently selected against, there also exists the possibility that spawning behaviours may play a role in limiting the number of hybrids produced. Although it is known that both morphs of smelt spawn in both streams at overlapping time, peak spawning periods between the two morphs may differ (P. Bentzen pers. obs.). Further investigation into this matter is needed. There is also the possibility of size-assortative mating as seen in sympatric sticklebacks (Gow *et al*., 2006). While results from my study and other similar ones suggest that reinforcement due to increased prezygotic isolation may be what prevent increased hybrid occurrences, further investigation into peak spawning times of both small and large smelt in Lochaber Lake will need to be taken into consideration.

The stability and the persistence of the sympatric smelt morphs depend on the continuous balance between hybridization and the selection against hybrids (Bradbury *et al*., 2011). There is

the possibility of a collapse of sympatric forms to a single population, something that may be often linked with environmental change and young post glacial lineages (Taylor *et al*., 2006; Bradbury *et al*., 2011). The collapse of Threespine Stickleback in Enos Lake, BC, occurred after the introduction of a crayfish species in the 1990's (Taylor *et al*., 2006). The stability of a sympatric pair can be compromised if the relative fitness of hybrids increases or if environmental cues that are associated with keeping populations diverted change (Bradbury *et al*., 2011). In Lochaber Lake, the observed genetic differentiation appears to be stable over the span of a decade, and moderate to high genetic differentiation is present, therefore selection appears to be efficient at removing hybrids (Bradbury *et al*., 2011). As noted in this study, hybrids are present in the Lochaber Lake population of smelt, therefore the future maintenance of the sympatric pair will be dependent on the relative fitness of the hybrids (Bradbury *et al*., 2011).

3.5 Conclusion

In summary, we examined whether or not larval smelt could be identified to morph based on their genetic makeup and if we could identify hybrid individuals in both the adult and larval stages of Rainbow smelt in Lochaber Lake. Based on our findings, the larvae examined were identified to morph in the same proportions that are observed in the adult smelt. We have also determined that hybrid individuals do exist in the adult and larval smelt, but that there exists some barrier to hybridization that acts toward them at a very early stage in their life cycle. Although prezygotic incompatibilities are likely responsible for the small number of hybrids, we would need to further investigate whether spawning behaviour is partially responsible for the low number of hybrids. More time observing the spawning behaviours between the small and the large morph of Rainbow smelt in Lochaber Lake would be needed.

CHAPTER 4 Thesis Conclusion

In summary, this study found that small and large smelt are genetically differentiated and that this genetic difference is stable over a several generations. I have also discovered that by using microsatellite loci we can identify larval smelt to morph. The power to detect hybrids beyond F1's was weak however, when adding q-values for all hybrid classes it was found that larval smelt and adult smelt had similar numbers of hybrids.

Appendices

Appendix 2.1. Population statistics for both morph, including the total number of fish sampled within each morph (N), total number of alleles (A), size range of alleles (R), expected (HE) and observed (HO) heterozygosity, inbreeding coefficient (FIS), mean number of alleles (NA), allelic richness (AE), Hardy-Weinberg equilibrium (PHWE).

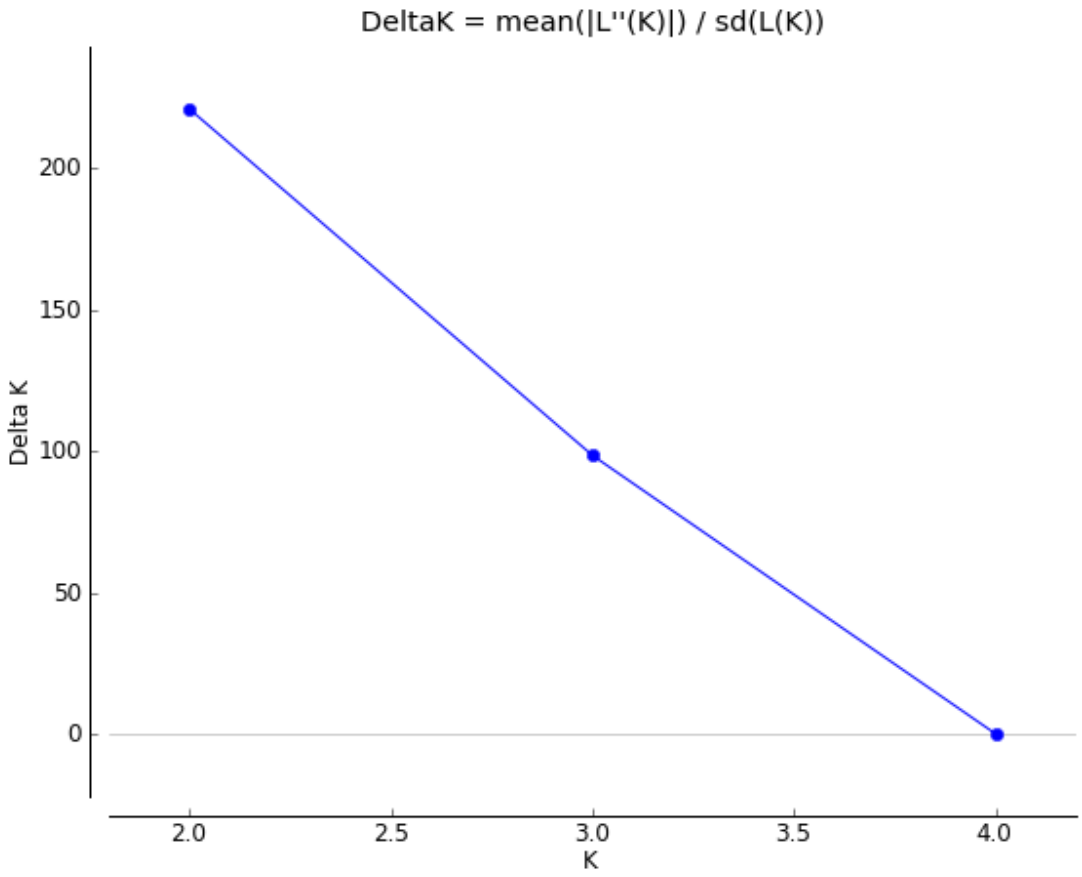
Ecotype		Locus						
		Omo1	**Omo2**	**Omo3**	**Omo4**	**Omo5**	**Omo11**	**Omo14**
	A	7	7	18	11	20	11	13
	R	108-166	172-196	114-242	168-212	222-298	157-197	120-248
	H_E	0.47344	0.68735	0.89262	0.73966	0.86989	0.76248	0.83371
Small N = 675	**H_O**	0.48137	0.60411	0.89724	0.74096	0.87768	0.73791	0.50755
	F_{IS}	-0.017	0.121	-0.005	-0.002	-0.009	0.032	0.391
	A_E	5.978	6.457	15.186	8.971	17.07	9.133	11.437
	PHWE	0.16128	0.00001	0.73158	0.45866	0.2601	0.01444	0
	A	3	5	14	6	9	5	8
	R	112-120	176-200	186-238	172-196	230-266	165-181	204-240
	H_E	0.3876	0.5111	0.88155	0.56542	0.54199	0.33014	0.75888
Large N = 236	**H_O**	0.43478	0.5414	0.88085	0.47034	0.5671	0.31304	0.61472
	F_{IS}	-0.122	-0.059	0.001	0.168	-0.046	0.052	0.19
	A_E	2.683	5	13.782	5.55	8.306	4.683	7.359
	PHWE	0.11974	0.01327	0.00073	0	0.0032	0.01613	0

Appendix 2.2. Linkage Disequilibrium and Hardy Weinberg equilibrium for both morph at each year sampled.

Linkage Disequilibrium		
Morph→	**Small**	**Large**
Year↓		
2002	No LD	No LD
2003	No LD	No LD
2004	Omo1&14	No LD
2010/2011	No LD	No LD

Hardy Weinberg Equilibrium		
Morph→	**Small**	**Large**
Year↓		
2002	Omo 14	All in HWE
2003	Omo14	Omo3,4,14
2004	Omo11&14	Omo3,4,14
2010/2011	Omo14	Omo14

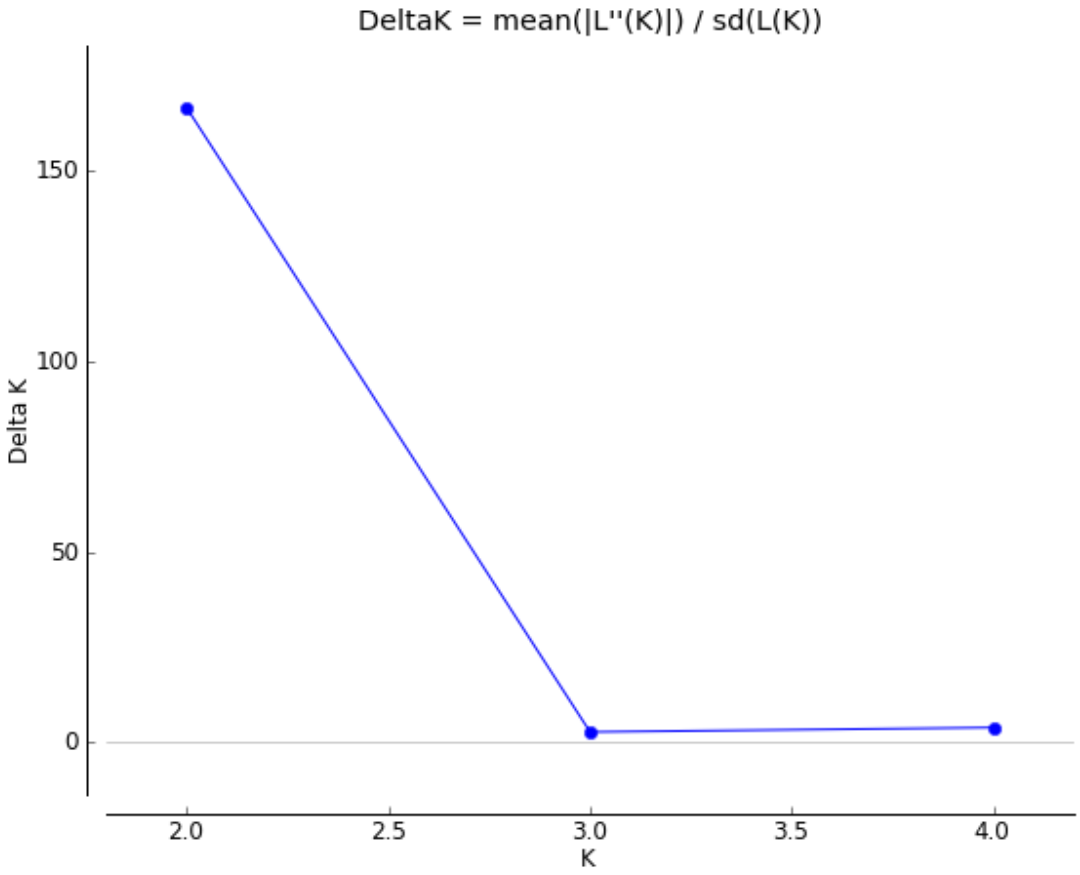
Appendix 2.3. Delta K plot for Bayesian clustering analysis using STRUCTURE of 4 populations of Rainbow Smelt in Lochaber Lake. K-values from K=2 to K=4 are shown, with one peak only at K=2.



Appendix 3.1. Population statistics for both morphs and larval smelt, including the total number of fish sampled within each morph (N), total number of alleles (A), size range of alleles (R), expected (H_E) and observed (H_O) heterozygosity, inbreeding coefficient (F_{IS}), mean number of alleles (NA), allelic richness (A_E) and Hardy-Weinberg equilibrium (PHWE).

		Locus								
		Omo1	**Omo2**	**Omo3**	**Omo4**	**Omo5**	**Omo9**	**Omo11**	**Omo14**	**Omo15**
	N	211	210	201	210	193	215	134	210	160
	A	5	6	17	9	17	15	9	17	9
	R	114-130	172-192	114-242	174-214	244-312	150-298	158-198	90-242	176-212
Small	H_E	0.455	0.636	0.891	0.747	0.871	0.669	0.746	0.872	0.734
N = 216	H_O	0.445	0.614	0.896	0.81	0.865	0.66	0.701	0.876	0.731
	F_{IS}	0.022	0.035	-0.005	-0.084	0.007	0.012	0.06	-0.005	0.004
	A_E	3.901	4.872	12.164	7.005	13.199	10.852	6.806	11.494	7.86
	PHWE	0.502	0.152	0.18	0.891	0.486	0.222	0.031	0.203	0.356
	N	54	56	58	58	55	58	34	56	44
	A	3	3	10	6	9	8	5	8	6
	R	114-122	176-184	190-230	174-198	248-284	150-202	166-182	138-238	176-196
Large	H_E	0.36	0.486	0.857	0.459	0.582	0.361	0.411	0.787	0.356
N = 58	H_O	0.426	0.5	0.948	0.483	0.545	0.397	0.265	0.714	0.318
	F_{IS}	-0.183	-0.029	-0.0106	-0.051	0.062	-0.098	0.356	0.093	0.105
	A_E	2.63	2.848	9.396	5.003	7.084	6.764	5	7.303	5.543
	PHWE	0.427	1	0.605	0.723	0.181	0.899	0.005	0.008	0.065
	N	308	303	299	297	260	313	278	314	306
	A	6	8	17	11	18	14	10	15	13
	R	114-134	172-200	182-246	170-218	240-308	146-198	158-194	134-318	176-208
Larvae	H_E	0.466	0.624	0.886	0.736	0.888	0.613	0.725	0.862	0.64
N = 348	H_O	0.552	0.65	0.886	0.603	0.808	0.604	0.683	0.834	0.565
	F_{IS}	-0.184	-0.041	-0.001	0.182	0.09	0.015	0.057	0.032	0.117
	A_E	4.13	5.188	12.439	7.269	13.567	9.572	7.403	10.66	7.807
	PHWE	0.01	0.951	0.768	0	0.022	0.693	0.003	0.276	0

Appendix 3.2. Delta K plot for Bayesian clustering analysis using STRUCTURE of 4 populations of Rainbow Smelt in Lochaber Lake. K-values from K=2 to K=4 are shown, with one peak only at K=2.



Appendix 3.3: ONCOR results depicting the probability of each individual larvae in the mixture belonging to either baseline population of small or large smelt.

	Small	Large
Larvae1	0.864665	0.135335
Larvae2	0.018542	0.981458
Larvae3	1	0
Larvae4	0.021903	0.978097
Larvae5	1	0
Larvae6	0.028148	0.971852
Larvae7	0.997881	0.002119
Larvae8	0.999999	0.000001
Larvae9	0.999639	0.000361
Larvae10	0.002392	0.997608
Larvae11	0.999624	0.000376
Larvae12	0.999868	0.000132
Larvae13	0.848746	0.151254
Larvae14	0.396777	0.603223
Larvae15	0.99999	0.00001
Larvae16	0.000679	0.999321
Larvae17	1	0
Larvae18	0.00006	0.99994
Larvae19	0.015574	0.984426
Larvae20	0.771743	0.228257
Larvae21	0.001027	0.998973
Larvae22	0.000099	0.999901
Larvae23	0.000191	0.999809
Larvae24	0.999996	0.000004
Larvae25	0.997416	0.002584
Larvae26	0.998366	0.001634
Larvae27	1	0
Larvae28	0.351188	0.648812
Larvae29	0.008221	0.991779
Larvae30	0.999902	0.000098
Larvae31	0.999997	0.000003
Larvae32	0.193252	0.806748
Larvae33	0.999986	0.000014
Larvae34	0.999954	0.000046
Larvae35	1	0
Larvae36	0.958873	0.041127
Larvae37	0.000992	0.999008
Larvae38	0.012407	0.987593

Larvae39	0.995824	0.004176
Larvae40	0.000197	0.999803
Larvae41	0.043438	0.956562
Larvae42	0.748928	0.251072
Larvae43	0.99997	0.00003
Larvae44	1	0
Larvae45	0.371285	0.628715
Larvae46	0.000061	0.999939
Larvae47	0.954135	0.045865
Larvae48	0.99992	0.00008
Larvae49	0.999975	0.000025
Larvae50	0.991257	0.008743
Larvae51	0.447203	0.552797
Larvae52	0.963574	0.036426
Larvae53	0.999636	0.000364
Larvae54	1	0
Larvae55	1	0
Larvae56	0.999873	0.000127
Larvae57	0.971386	0.028614
Larvae58	0.999936	0.000064
Larvae59	0.979982	0.020018
Larvae60	1	0
Larvae61	1	0
Larvae62	0.999253	0.000747
Larvae63	0.999994	0.000006
Larvae64	0.993436	0.006564
Larvae65	1	0
Larvae66	1	0
Larvae67	0.999934	0.000066
Larvae68	0.999993	0.000007
Larvae69	0.261131	0.738869
Larvae70	0.99981	0.00019
Larvae71	0.999358	0.000642
Larvae72	0.999891	0.000109
Larvae73	0.996955	0.003045
Larvae74	1	0
Larvae75	0.999789	0.000211
Larvae76	1	0
Larvae77	0.999994	0.000006
Larvae78	0.999989	0.000011
Larvae79	0.999979	0.000021
Larvae80	1	0
Larvae81	0.934504	0.065496

Larvae82	0.405282	0.594718
Larvae83	0.999778	0.000222
Larvae84	0.999531	0.000469
Larvae85	0.999999	0.000001
Larvae86	1	0
Larvae87	0.957275	0.042725
Larvae88	1	0
Larvae89	0.999723	0.000277
Larvae90	1	0
Larvae91	0.999671	0.000329
Larvae92	1	0
Larvae93	0.999649	0.000351
Larvae94	0.999994	0.000006
Larvae95	1	0
Larvae96	0.999996	0.000004
Larvae97	0.999996	0.000004
Larvae98	1	0
Larvae99	0.999991	0.000009
Larvae100	1	0
Larvae101	0.999999	0.000001
Larvae102	1	0
Larvae103	0.994238	0.005762
Larvae104	0.999991	0.000009
Larvae105	1	0
Larvae106	0.99996	0.00004
Larvae107	0.999996	0.000004
Larvae108	0.999996	0.000004
Larvae109	1	0
Larvae110	0.858696	0.141304
Larvae111	0.20515	0.79485
Larvae112	0.000011	0.999989
Larvae113	1	0
Larvae114	1	0
Larvae115	0.001793	0.998207
Larvae116	0.998552	0.001448
Larvae117	0.999962	0.000038
Larvae118	0.990554	0.009446
Larvae119	0.999141	0.000859
Larvae120	0.000119	0.999881
Larvae121	0.005939	0.994061
Larvae122	0.999997	0.000003
Larvae123	0.972075	0.027925
Larvae124	0.000105	0.999895

Larvae125	0.99996	0.00004
Larvae126	0.864321	0.135679
Larvae127	0.996007	0.003993
Larvae128	0.57979	0.42021
Larvae129	0.999252	0.000748
Larvae130	0.999444	0.000556
Larvae131	0.996872	0.003128
Larvae132	0.999813	0.000187
Larvae133	0.967818	0.032182
Larvae134	1	0
Larvae135	0.976128	0.023872
Larvae136	0.999994	0.000006
Larvae137	0.776548	0.223452
Larvae138	0.997537	0.002463
Larvae139	0.999999	0.000001
Larvae140	0.790856	0.209144
Larvae141	0.999221	0.000779
Larvae142	0.999989	0.000011
Larvae143	0.999967	0.000033
Larvae144	0.999004	0.000996
Larvae145	0.999916	0.000084
Larvae146	1	0
Larvae147	1	0
Larvae148	0.789189	0.210811
Larvae149	1	0
Larvae150	0.999673	0.000327
Larvae151	0.954821	0.045179
Larvae152	0.998973	0.001027
Larvae153	0.000128	0.999872
Larvae154	0.994657	0.005343
Larvae155	0.999181	0.000819
Larvae156	1	0
Larvae157	0.999999	0.000001
Larvae158	0.998763	0.001237
Larvae159	0.995754	0.004246
Larvae160	0.985521	0.014479
Larvae161	0.602538	0.397462
Larvae162	0.331663	0.668337
Larvae163	0.942853	0.057147
Larvae164	0.500159	0.499841
Larvae165	0.999884	0.000116
Larvae166	0.000241	0.999759
Larvae167	0.994852	0.005148

Larvae168	1	0
Larvae169	1	0
Larvae170	0.999999	0.000001
Larvae171	0.04297	0.95703
Larvae172	1	0
Larvae173	0.997984	0.002016
Larvae174	0.000003	0.999997
Larvae175	0.999949	0.000051
Larvae176	0.999999	0.000001
Larvae177	0.996315	0.003685
Larvae178	0.000369	0.999631
Larvae179	0.004222	0.995778
Larvae180	0.999757	0.000243
Larvae181	0.999966	0.000034
Larvae182	0.998643	0.001357
Larvae183	0.999999	0.000001
Larvae184	0.999999	0.000001
Larvae185	0.999484	0.000516
Larvae186	1	0
Larvae187	0.030746	0.969254
Larvae188	0.999957	0.000043
Larvae189	0.999223	0.000777
Larvae190	1	0
Larvae191	0.999999	0.000001
Larvae192	0.997398	0.002602
Larvae193	1	0
Larvae194	0.000343	0.999657
Larvae195	0.103555	0.896445
Larvae196	0.999934	0.000066
Larvae197	1	0
Larvae198	1	0
Larvae199	0.996615	0.003385
Larvae200	0.000065	0.999935
Larvae201	1	0
Larvae202	0.518875	0.481125
Larvae203	0.992323	0.007677
Larvae204	1	0
Larvae205	0.999998	0.000002
Larvae206	0.875731	0.124269
Larvae207	0.002865	0.997135
Larvae208	0.987965	0.012035
Larvae209	1	0
Larvae210	1	0

Larvae211	0.996871	0.003129
Larvae212	0.566988	0.433012
Larvae213	0.212237	0.787763
Larvae214	0.89774	0.10226
Larvae215	0.447661	0.552339
Larvae216	0.999996	0.000004
Larvae217	1	0
Larvae218	1	0
Larvae219	1	0
Larvae220	1	0
Larvae221	0.996636	0.003364
Larvae222	0.000037	0.999963
Larvae223	0.999303	0.000697
Larvae224	0.016832	0.983168
Larvae225	0.999996	0.000004
Larvae226	0.988344	0.011656
Larvae227	0.959052	0.040948
Larvae228	0.058535	0.941465
Larvae229	0.910453	0.089547
Larvae230	1	0
Larvae231	0.999915	0.000085
Larvae232	1	0
Larvae233	0.999998	0.000002
Larvae234	0.739712	0.260288
Larvae235	1	0
Larvae236	0.213468	0.786532
Larvae237	1	0
Larvae238	0.189203	0.810797
Larvae239	0.999999	0.000001
Larvae240	0.998671	0.001329
Larvae241	0.004729	0.995271
Larvae242	0.000314	0.999686
Larvae243	1	0
Larvae244	0.999081	0.000919
Larvae245	0.067487	0.932513
Larvae246	0.009673	0.990327
Larvae247	1	0
Larvae248	0.999945	0.000055
Larvae249	1	0
Larvae250	0.999558	0.000442
Larvae251	1	0
Larvae252	1	0
Larvae253	0.999996	0.000004

Larvae254	0.999974	0.000026
Larvae255	1	0
Larvae256	0.999584	0.000416
Larvae257	0.328048	0.671952
Larvae258	0.000008	0.999992
Larvae259	0.983101	0.016899
Larvae260	0.004391	0.995609
Larvae261	0.997853	0.002147
Larvae262	0.999999	0.000001
Larvae263	0.999699	0.000301
Larvae264	0.999998	0.000002
Larvae265	0.999993	0.000007
Larvae266	0.999996	0.000004
Larvae267	0.999999	0.000001
Larvae268	1	0
Larvae269	1	0
Larvae270	0.991449	0.008551
Larvae271	0.999992	0.000008
Larvae272	0.000222	0.999778
Larvae273	0.999903	0.000097
Larvae274	1	0
Larvae275	0.999718	0.000282
Larvae276	0.999999	0.000001
Larvae277	0.00495	0.99505
Larvae278	0.00153	0.99847
Larvae279	0.99952	0.00048
Larvae280	0.999934	0.000066
Larvae281	0.999879	0.000121
Larvae282	0.999396	0.000604
Larvae283	0.140671	0.859329
Larvae284	1	0
Larvae285	0.999995	0.000005
Larvae286	1	0
Larvae287	0.999994	0.000006
Larvae288	0.999403	0.000597
Larvae289	0.999976	0.000024
Larvae290	0.98105	0.01895
Larvae291	0.978052	0.021948
Larvae292	1	0
Larvae293	1	0
Larvae294	0.985326	0.014674
Larvae295	1	0
Larvae296	1	0

Larvae297	1	0
Larvae298	0.999997	0.000003
Larvae299	1	0
Larvae300	0.999573	0.000427
Larvae301	1	0
Larvae302	0.999198	0.000802
Larvae303	0.999993	0.000007
Larvae304	0.999984	0.000016
Larvae305	0.011584	0.988416
Larvae306	0.992638	0.007362
Larvae307	1	0
Larvae308	1	0
Larvae309	0.999946	0.000054
Larvae310	1	0
Larvae311	1	0
Larvae312	1	0
Larvae313	0.999989	0.000011
Larvae314	0.98038	0.01962
Larvae315	0.999988	0.000012
Larvae316	0.999998	0.000002
Larvae317	1	0
Larvae318	1	0
Larvae319	0.999988	0.000012
Larvae320	0.134452	0.865548
Larvae321	0.999975	0.000025
Larvae322	1	0
Larvae323	0.999817	0.000183
Larvae324	0.974601	0.025399
Larvae325	1	0
Larvae326	0.99999	0.00001
Larvae327	1	0
Larvae328	1	0
Larvae329	0.993422	0.006578
Larvae330	0.999987	0.000013
Larvae331	0.999723	0.000277
Larvae332	0.999939	0.000061
Larvae333	1	0
Larvae334	0.99779	0.00221
Larvae335	1	0
Larvae336	0.993035	0.006965
Larvae337	0.999858	0.000142
Larvae338	0.999988	0.000012
Larvae339	0.99988	0.00012

Larvae340	0.999999	0.000001
Larvae341	0.533608	0.466392
Larvae342	0.999727	0.000273
Larvae343	0.999631	0.000369
Larvae344	0.793005	0.206995
Larvae345	1	0
Larvae346	0.004268	0.995732
Larvae347	1	0
Larvae348	0.999962	0.000038

BIBLIOGRAPHY

Anderson EC, Thompson EA. 2002. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics.* **160:** 1217-1229

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*, Salmonidea) populations from microsatellites. *Molecular Biology and Evolution* **15**: 143-159

Baby MC, Bernatchez L, Dodson JJ. 1991. Genetic structure and relationships among anadromous and landlocked populations of rainbow smelt, *Osmerus mordax*, Mitchill, as revealed by mtDNA restriction analysis. *Journal of Fish Biology* **39**: 61-68

Baillie SM, Muir AM, Scribner K, Bentzen P, Krueger CC. 2016. Loss of genetic diversity and reduction of genetic distance among lake trout *Salvelinus namaycush* ectomorphs, Lake Superior 1959 to 2013. *Journal of Great Lakes Research* **42:** 204-216

Ballard JWO, Kreitman M. 1995. Is mitochondrial DNA a strictly neutral marker? *Trends in Ecology & Evolution* **10**: 485-488

Balloux F, Lugon-Moulin N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**: 155-165

Banks MA, Eichert W, Olsen JB. 2003. Which genetic loci have greater population assignment power? *Bioinformatics*. **19:** 1436-1438

Barrette MF, Daigle G, Dodson JJ. 2009. Intraspecific vicariant history and the evolution of adaptive morphological diversity within species. *Biological Journal of the Linnaean Society* **97:** 140-151

Barton NH. 2001. Speciation. *TRENDS in Ecology & Evolution* **16:** 325

Behnke RJ. 1972. The systematic of salmondid fishes of recently glaciated lakes. *Journal of the Fisheries Research Board of Canada* **29**: 639-671

Bell M. 1976. Reproductive character displacement in threespine sticklebacks. *Evolution*. **30**: 847-848

Bell M. 1989. Morphogenesis and evolution. *Quarterly Review of Biology.* **64**: 345-346

Bentzen P, Ridgway MS, McPhail JD. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): spatial segregation and seasonal habitat shifts in the Enos Lake species pair. *Canadian Journal of Zoology* **62**: 2436-2439

Bernatchez L. 1997. Mitochondrial DNA analysis confirms the existence of two glacial races of rainbow smelt *Osmerus mordax* and their reproductive isolation in the St Lawrence River estuary (Quebec, Canada) *Molecular Ecology* **6**: 73-83

Bernatchez L, Dodson JJ. 1990. Allopatric origin of sympatric populations of Lake Whitefish (*Coregonus clupeaformis*) revealed by mitochondrial DNA restriction analysis. *Evolution* **44**: 1263-1271

Bernatchez L, Dodson JJ. 1991. Phylogeographic structure in mitochondrial DNA of Lake Whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. *Evolution* **45**: 1016-1035

Bernatchez L, Martin S. 1996. Mitochondrial DNA diversity in anadromous rainbow smelt, *Osmerus mordax* Mitchill: a genetic assessment of the member-vagrant hypothesis. *Canadian Journal of Fisheries and Aquatic Sciences* **53**: 424-433

Bernatchez L, Chouinard A, Lu G. 1999. Intergrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society* **68:** 173-194

Birky CW, Maruyama T, Fuerst PA. 1983. An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics* **103**: 513-527

Blackie C, Weese D, Noakes DLG. 2003. Evidence for resource polymorphism in the lake charr (*Salvelinus namaycush*) population of Great Bear Lake, Northwest Territories, Canada. *Écoscience* **10**: 410-415

Bradbury IR, Coulson MW, Campana SE, Bentzen P. 2006. Morphological and genetic differentiation in anadromous smelt *Osmerus mordax* (Mitchill): disentangling the effects of geography and morphology on gene flow. *Journal of Fish Biology* **69**: 95-114

Bradbury IR, Coulson MW, Cook AM, Bentzen P. 2010. Ecidence for divergence and adaptive isolation in post-glacially derived bimodal allopatric and sympatric rainbow smelt populations. *Biological Journal of the Linnean Society* **101**: 583-594

Bradbury IR, Bradford R, Bentzen P. 2011. Genetic and phenotypic diversity and divergence in sympatric Lake Utopia rainbow smelt (*Osmerus mordax*). *DFO Canadian Science Advisory Secretariat Research Document.* 2011/008: vi + 28p.

Brown WM, George MJ, Wilson AC. 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences USA* **76**: 1967-1971

Buckley JL. 1989. Species profile: life histories and environmental requirements of coastal fishes and invertebrates (North Atlantic)—rainbow smelt. *U.S. Fisheries and Wildlife Service Biology*, Rep. 82(11.106). U.S Army Corps Engineers, TR EL-82-4. 11pp

Carlson SM, Rich HB, Quinn TP. 2009. Does variation in selection imposed by bears drive divergence among populations in the size and shape of sockeye salmon? *Evolution* **63**: 1244-1261

Chouinard A, Pigean D, Bernatchez L. 1996. Lack of specialization in trophic morphology between genetically differentiated dwarf and normal forms of lake whitefish (*Coregonus clupeaformsi*) in Lac de l'Est, Quebec. *Canadian Journal of Zoology* **74**: 1989-1998

Claytor RR, Vespoor E. 1991. Discordant phenotypic variation in sympatric resident and anadromous Atlantic salmon (*Salmo salar*) populations. *Canadian Journal of Zoology* **69**: 2846-2852

Copeman DG. 1977. Population differences in rainbow smelt, *Osmerus mordax*: multivariate analysis of mensural and meristic data. *Journal of the Fisheries Research Board of Canada* **34**: 1220-1229

Copeman DG, McAllister DE. 1978. Analysis of the effect of transplantation on morphometric and meristic characters in lake populations of the rainbow smelt, *Osmerus mordax* (Mitchill). *Enbironmental Biology of Fishes* **3**: 253-259

Coulson MW, Paterson IG, Green A, Kepkay R, Bentzen P. 2006. Characterization of di- and tetranucleotide microsatellite markers in rainbow smelt (*Osmerus mordax*). *Molecular Ecology Notes.* **6:** 942-944

Coyne JA. 1992. Genetics and speciation. *Nature* **355**: 511-515

Coyne JA, Orr HA. 2004. Speciation. Sinauer Associates, Inc., Sunderland, MA.

Cracraft J. 1983. Species concepts and speciation analysis. *Current Ornithology* **1:** 159-187

Deiringer D, Schlotterer C. 2003. MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3:** 167-169

Dodd DMB. 1989. Reproductive isolation as a consequence of adaptive divergence in *Drosophila pseudoobscura*. *Evolution*. **43**: 1308-1311

Doebeli M, Dieckmann U. 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. *The American Naturalist* **156**: s77-2101

Dowling TE, Moor WS. 1985. Genetic variation and divergence of the sibling pare of cyprinid fishes, *Notropis cornutus* an *N. chrysocephalus. Biochemical Systematics and Biology.* **13:** 471-476

Earl D, von Holdt BM. 2011. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*. **4:** 359-361

Elphinstone MS, Hinten GN, Anderson MJ, Nock CJ. 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. *Molecular Ecology.* **3:** 317-320

Estoup A, Angers B. 1998. Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In: *Advances in Molecular Ecology* (ED. By Carbalho GR), pp. 55-86. IOS Press, Amsterdam.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*. **14**: 2611-2620

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform populations genetics analyses under Linux and Windows. *Molecular Ecology Resources.* **10**: 564-567

Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB. 2011. Extent and scale of local adaptation in salmonid fishes: Review and meta-analysis. *Heredity* **106**: 404-420

Frost WE. 1965. Breeding habits of Windermere charr, *Salvelinus willughbii* (Gunther), and their bearing on speciation of these fishes. *Proceedings of the Royal Society B: Biological Sciences* **163:** 232-284

Fendersen OC. 1964. Evidence of subpopulations of lake whitefish, *Coregonus clupeaformis*, involving a dwarf form. *Transactions of the American Fisheries Society* **93**: 77-94

Funk DJ, Nosil P. 2008. Comparative analyses of ecological speciation. In: Tilmon K, ed. *Specialization, speciation and radiation: the evolutionary biology of herbivorous insects.* Berkeley, CA: University of California Press, 117-135

Gavrilets S, Vose A. 2005. Dynamic patterns of adaptive radiation. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 18040-18045

Goudet J. 1995. Fstat Version 1.2: A computer program to calculate F-statistics. *Journal of Heredity.* **86**: 485

Goudet J, Raymond M, de Meeus T, Rousset F. 1996. Testing differentiation in diploid populations. *Genetics* **139**: 463-471

Gow JL, Peichel CL, Taylor EB. 2006. Contrasting hybridization rates between sympatric three-spined sticklebacks highlight the fragility of reproductive barriers between evolutionarily young species. *Molecular Ecology*

Gow JL, Peichel CL, Taylor EB. 2007. Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. *Journal of Evolutionary Biology* **20**: 2173-2180

Grant PR, Grand BR, Markert JA, Keller LF, Petren K. 2004. Convergent evolution of Darwin's finches caused by introgressive hybridization and selection. *Evolution* **58**: 1588-1599

Harris LN, Chavarie L, Bajno R, Howland KL, Wiley SH, Tonn WM, Taylor EB. 2015. Evolution and origin of sympatric shallow-water morphotypes of Lake Trout, *Salvelinus namaycush*, in Canada's Great Bear Lake. *Heredity* **114**: 94-106

Hatfield T, Schluter D. 1996. A test for sexual selection on hybrids of two sympatric sticklebacks. *Evolution* **50**: 2429-2434

Hatfield T, Schluter D. 1999. Ecological speciation in sticklebacks: Environment-dependent hybrid fitness. *Evolution* **53**: 1491-1505

Hull DL. 1997. The ideal species concept – and why we can't get it. In: Claridge, M.F., Dawah, H.A. and Wilson, M.R. Eds: *Species: The Units of Biodiversity*. Chapman and Hall, New York, pp. 357-380

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

Kalinowski ST, Manlove KR, Taper ML. 2007. ONCOR: A computer program for genetic stock identification. http://www.montana.edu/kalinowski/Software/ONCOR.htm

Lanteigne J, McAllister DE. 1983. The pygmy smelt, *Osmerus spectrum* Cope, 1870, a forgotten sibling species of Eastern North American Fish. Ichthyology Section, National Museum of Natural Sciences, Canada.

Levin, DA. 1979. The nature of plant species. *Science* **204**: 381-384

Lowry BD, Modliszewski JL, Wright KM, Wu CA, Willis JH. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society of Biology* **363:** 3009-3021

Lu G, Bernatchez L. 1998. Experimental evidence for reduced hybrid viability between dwarf and normal ecotypes of lake whitefish (*Coregonus clupeaformis* Mitchill). *Proceedings of the Royal Society of London Series B, Biological Sciences* **265:** 1025-1030

Lu G, Bernatchez L. 1999. Correlated trophic specialization and genetic divergence in sympatric Lake Whitefish ecotypes (*Coregonus clupeaformis*): Support for the ecological speciation hypothesis. *Evolution* **53**: 1491-1505

May-McNally SL, Quinn TP, Taylor EB. 2015. Low levels of hybridization between sympatric Arctic char (*Salvelinus alpinus*) and Dolly Varden char (*Salvelinus malma*) highlights their genetic distinctiveness and ecological segregation. *Ecology and Evolution.* **5:** 3031-3045

Mayden RL. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. In: Claridge, M.F., Dawah, H.A. and Wilson, M.R. Eds: *Species: The Units of Biodiversity*. Chapman and Hall, New York, pp. 381-424

Mayr, E. 1940. Speciation phenomena in birds. *American Naturalist* **74**: 249-278

Mayr, E. 1963. *Animal Species and Evolution*. The Balknap Press of Harvard University, Cambridge, MA. 974 pp.

McPhail JD. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic evidence for a species pair in Enos Lake. *Canadian Journal of Zoology*. **62**: 1402-1408

McPhail JD. 1992. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic evidence for a species pair in Paxton Lake. *Canadian Journal of Zoology* **70**: 361-369

McPhail JD. 1993. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): origin of the species pairs. *Canadian Journal of Zoology* **71**: 515-523

McPhail JD. 1994. Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south western British Columbia. In: Bell MA, Foster SA, eds *The evolutionary biology of the threespine stickleback*. Oxford: Oxford University Press, 399-437

Minamory S. 1956. Physiological isolationg in Cobitidae. IV. Speciation of two sympatric races of Lake Biwa of the striated spinous loach. *Japanese Journal of Zoology* **12**: 89-104

Murawski, SA, Clayton G, Reed RJ, Cole CF. 1980. Movement of spawning Rainbow smelt, *Osmerus mordax*, in a Mssachusetts estuary. *Estuaries* **4**: 308-314

Nei M. 1987. *Molecular evolutionary genetics* Columbia University Press, USA.

Nellbring S. 1989. The ecology of smelts (genus *Osmerus*): a literature review. *Nordic Journal of Freshwater Research*, **65:** 116-145

Nielsen EE, Bach LA, Kotlicki P. 2006. HYBRIDLAB (version 1.0): a program for generating simulated hybrids from population samples. *Molecular Ecology Notes*. **6:** 971-973

Nosil PT, Vines H, Funk DJ. 2005. Reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **59:** 705-719

Nosil P. 2007. Divergent host plant adaptation and reproductive isolation between ecotypes of

Timema cristinae walking sticks. *American Naturalist* **169:** 151-162

Okamoto H. 1989. A genetic comparison of sympatric populations of sand lance (Genus *Ammodytes*) from the region east of Cape Soya, Japan. *Canadian Journal of Fisheries and Aquatic Sciences* **46**: 1945-1951

Osinov AG, Senchukova AL, Mugue NS, Pavlov SD, Chereshnev IA. 2015. Speciation ang genetic divergence of three species of charr from ancient Lake El'gygytgyn (Chukotka) ad their phylogenic relationships with other representatives of the genus *Salvelinus*. *Biological Journal of the Linnean Society* **116:** 63-85

Otte D, Endler JA. (*eds*) **1989.** *Speciation and its consequences.* Sunderland, MA: Sinauer

Palumbi SR, Baker CS. 1994. Contrasting population structure from nuclear intron sequence and mtDNA of humpback whales. *Molecular Biology and Evolution* **11**: 426-435

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

Planes S, Doherty PJ. 1997. Genetic and color interations at a contact zone of *Acanthochromis polyacanthus*: a marine fish lacking pelagic larvae. *Evolution* **51**: 1232-1243

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

Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences of the United States of America.* **94:** 9197-9201

Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223-225

Robinson B, Wilson D. 1994. Character release and displacement in fishes: a neglected literature. *American Naturalist* **144**: 596-627

Robinson BW, Schluter D. 2000. Natural selection and the evolution of adaptive genetic variation in norther freshwater fishes. In *Adaptive Genetic Variation in the Wild* (Mousseau, T. A. Sinervo, B. & Endler, J. A., eds), pp. 65-94, New York: Oxford University Press.

Rogers S, Bernatchez L. 2006. The genetic basis of intrinsic and extrinsic postzygotic isolation jointly promoting speciation in the lake whitefish species complex (*Coregonus clupeaformis*). *Journal of Evolutionary Biology* **19**: 1979-1994

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes*. **4**: 137-138

Rueffler C, Van Dooren TJM, Leimar L, Abrams PA. 2006. Disruptive selection and then

what? *TRENDS in Ecology and Evolution* **21**: 238-245

Rundle HD. 2002. A test of ecologically dependent postmating isolation between sympatric sticklebacks. *Evolution* **56**: 233-329

Rundle HD, Nosil P. 2005. Ecological speciation. *Ecology Letters* **8**: 336-352

Rundle HD, Schluter D. 1998. Reinforcement of stickleback mate preferences: sympatry breed contempt. *Evolution* **52**: 200-208

Rupp RS. 1968. Life history and ecology of the smelt (*Osmerus mordax*) in the inland waters of Maine. Final Report, Federal Aid in Sport Fish Restoration Act Project. F-10-R, Maine Dep. Inland Fish Game. 36pp

Ruzzante DE, Walde SJ, Cussac VE, Macchi PJ, Alonso MF. 1998. Trophic polymorphism, habitate and diet segregation in *Percichthys trucha* (Pisces: Percichthyidae) in the Andes. *Biological Journal of the Linnean Society* **65**: 191-214

Saint-Laurent R, Legault M, Bernatchez L. 2003. Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). *Molecular Ecology* **12:** 315-330

Schluter D. 1995. Adaptive radiation in sticklebacks: Trade-offs in feeding performance and growth. *Ecology* **76**: 82–90.

Schluter, D. **1996a.** Ecological causes of adaptive radiation. American Naturalist **148** (Supplement): S40–S64.

Schluter, D. 1996b. Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society of London*, **351**: 807–814.

Schluter D. 2000. Ecological character displacement in adaptive radiation. *American Naturalist* **156** (Suppl.): S4-S16

Schluter D. 2001. Ecology and the origin of species. **16**: 372-380

Schluter D. 2009. Evidence for ecological speciation and its alternative. *Science*. **323:** 737-741

Schluter D. 2010. The ecology of adaptive radiation. OUP: Oxford, 296pp.

Schluter D, McPhail J. 1992. Ecological character displacement and speciation in sticklebacks. *American Naturalist*. **140**: 85-108

Schluter D, Conte GL. 2009. Genetics and ecological speciation. *Procedings of the National Academy of Sciences of the United States of America* **106:** 9955-9962

Schluter D, McPhail J. 1992. Ecological character displacement and speciation in sticklebacks. *America Naturalist*. **140**: 85-108

Skulason S, Smith TB. 1995. Resource polymorphism in vertebrates. *TRENDS in Ecology and Evolution* **10**: 366-370

Smouse PE, Peakall R. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity*. **82**: 561-573

Smith, TB. Skúlason S. 1996. Evolutionary significance of resource polymorphism in fishes, amphibians, and birds. *Annual Review of Ecology and Systematics,* **27**: 111-133

Snorrason SS, Skúlason S, Jonsoon B, Malmquist HJ, Jónasson PM, Sanlund OT, Lindem T. 1994. Trophic specialization in Arctic charr *Salvelinus alpines* (Pisces; Salmonidae): morphological divergence and ontogenetic niche shifts. *Biological Journal of the Linnean Society* **52**: 1-18

Taylor EB. 1999. Species pairs of north temperate freshwater fishes: Evolution, taxonomy, and conservation. *Reviews in Fish Biology and Fisheries* **9:** 299-324

Taylor EB, Bentzen P. 1993a. Molecular genetic evidence for reproductive isolation between sympatric populations of smelt *Osmerus* in Lake Utopia, south-western New Brunswick, Canada. *Molecular Ecology* **2:** 345-357

Taylor EB, Bentzen P. 1993b. Evidence for multiple origins and sympatric divergence of trophic ecotypes of smelt (*Osmerus*) in northeastern North America. *Evolution.* **47:** 813-832

Taylor EB, McPhail J. 1999. Evolutionary history of an adaptive radiation in species pairs of threespine sticklebacks (*Gasterosteus*): insights from mitochondrial DNA. *Biological Journal of the Linnean Society* **66**: 271-291

Taylor EB, McPhail J. 2000. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus. Proceedings of the Royal Society of London, Series B-Biological Sciences* **267**: 2375-2384

Taylor E, Boughman J, Groenenboom M, Schluter M, Sniatynski M, Gow J. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of threespine stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* **15**: 343-355

Taylor EB, Gerlinsky C, Farrell N, Gow JL. 2011. A test of hybrid growth disadvantage in wild, free-ranging species pairs of Threespine Stickleback (*Gasterosteaus aculeatus*) and its implications for ecological speciation. *Ecolution*. **66**: 240-251

Tessier N, Bernatchez L. 1999. Stability of population structure and genetic diversity across generations assessed by microsatellites among sympatric populations of landlocked Atlantic Salmon (*Salmo salar* L.). *Molecular Ecology* **8**: 169-179

Turner BJ, Grosse DJ. 1980. Trophic differentiation in *Ilyodon*, a genus of stream-dwelling goodeid fishes: Speciation versus trophic polymorphism. *Evolution* **34**: 259-270

Väli Ü, Saag P, Dumbrovski V, Bernd-Ulrich M, Grzegorz M, Mizera T, Treinys R, Fagerberg S. 2010. Microsatellites and single nucleotide polymorphisms in avian hybrid identification: a comparative case study. *Journal of Avian Biology* **41**: 34-49

Vamosi SM, Hatfield T, Schluter D. 2000. A test of ecological selection against young-of-the-year hybrids of sympatrick sticklebacks. *Journal of Fish Biology* **57**: 109-121

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. **2004.** MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535-538.

Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*. **163**: 1177-1191

Wirtz P. 1999. Mother species-father species: unidirectional hybridization in animals with female choice. *Animal behavior*. **58:** 1-12