Spatial Distribution of Neutral Genetic Variation in a Wide Ranging Anadromous Clupeid, the American Shad (*Alosa sapidissima*)

by

Daniel J. Hasselman

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at

Dalhousie University

Halifax, Nova Scotia

May 2010

DALHOUSIE UNIVERSITY

DEPARTMENT OF BIOLOGY

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "Spatial Distribution of Neutral Genetic Variation in a Wide Ranging Anadromous Clupeid, the American Shad (*Alosa sapidissima*)" by Daniel J. Hasselman in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dated: May 26, 2010

External Examiner: _____

Research Supervisor: _____

Examining Committee: _____

_____

Departmental Representative: _____

DALHOUSIE UNIVERSITY

DATE: May 26, 2010

AUTHOR: Daniel J. Hasselman

TITLE: Spatial Distribution of Neutral Genetic Variation in a Wide Ranging Anadromous Clupeid, the American Shad (*Alosa sapidissima*)

DEPARTMENT OR SCHOOL: Department of Biology

DEGREE: PhD CONVOCATION: October YEAR: 2010

This thesis is dedicated to my parents. They have always offered unwavering support and words of encouragement throughout the course of my studies, and in all of my pursuits.

To my father. Over the course of my life you have continually demonstrated that there is no substitute for hard work. Some doubted my abilities to complete this degree and claimed that I was 'wasting my time'. However, you have instilled in me the ability to dream big, and to follow my heart no matter where it leads. Most importantly you have time and again shown that if you want something in life you go get it, and to hell with whatever anyone else thinks. Thank you for your life lessons, they have always served me well.

To my mother. There were plenty of times during the course of my studies where I wanted to throw up my hands, and walk away from this project entirely. However, that would have been the easy way out. Your personal struggles have served as an inspiration throughout the course of my research. No matter how many setbacks you have encountered over the last seven years you have never given up. In light of your determination to battle onwards against all odds, how could I possibly have justified giving up on my dreams? You are the most courageous person I have ever known, and I love you for all that you are. Keep fighting. NEVER give up!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Species long-term persistence is to a great extent contingent on the ability of populations to mount variable responses to perturbations; the breadths of which are largely dependent on the amount of heritable variation present at the population level. However, populations are not necessarily equivalent in their amount of genetic variation, or in their responses to future environmental conditions, and information about the magnitude and spatial distribution of intraspecific genetic variation is integral to conservation planning, and preserving species evolutionary potential. Using neutral molecular markers, I demonstrate that latitude is an important determinant of the amount and spatial distribution of genetic variation within and among Nearctic fishes. Latitudinal declines observed among species were mirrored by declines within species, and encourage a cautionary approach for interspecific comparisons and inferences of broad spatial patterns of genetic variation when data for individual species are obtained from only a portion of their range. I subsequently examined the magnitude and spatial distribution of microsatellite based variation for 33 spawning runs from across the range of American shad (*Alosa sapidissima*). Sequential reductions of intraspecific genetic variation with latitude were observed among spawning runs from formerly glaciated regions; consistent with stepwise post-glacial range expansion, and successive population founder events. Canadian populations exhibited temporally stable genetic differentiation characterized by a significant pattern of isolation by distance, and exhibited evidence of metapopulation structure. Although isolation by distance was observed among U.S. spawning runs, population structure was comparatively weak; a possible consequence of increased levels of gene flow (human-mediated or natural). Different spatial patterns of population structure were detected across the species' range, but are not likely due to alternative reproductive strategies (iteroparity vs. semelparity). Rather, these different patterns probably reflect different management strategies in Canada (no stocking) and the United States (stocking), alternative glacial histories, or combinations thereof. Reciprocal patterns of genetic diversity and differentiation observed across the species' range suggests that U.S spawning runs contribute more to diversity and less to differentiation than Canadian spawning runs. These results have implications for future shad restoration efforts, and the long-term persistence of the species.

ACKNOWLEDGEMENTS

It is impossible to conduct a study of this magnitude without the guidance of a team of supervisors, an army of volunteers, and the support of family and friends. I hope that I have remembered everyone, but if your name does not appear below it was either to maintain your privacy (i.e. recreational/commercial fishers), or because my short-term memory is failing me - my sincerest apologies if your name does not appear below but should be.

Paul Bentzen - I couldn't have asked for a better supervisor, mentor, colleague, and friend. When I first came to you six years ago I had no previous experience in molecular ecology, but you didn't hesitate in providing me with an opportunity to learn a biological sub-discipline that I had always found intimidating. Through all of the challenges I have faced over the course of this project (theoretical, methodological, analytical) your door has always been open, and you've always been there to answer my questions, no matter how mundane. This project could not have reached fruition without your guidance. Thank you for your incredible support and patience.

Rod Bradford - I owe you a debt of gratitude that words can hardly express. Your guidance over the course of two theses has largely defined my way of thinking, and the naïve undergraduate student you met so many years ago seems a distant memory. Your challenge of me to think critically and to continually strive to 'raise the bar' has played a highly significant ($p<0.001$; χ^2) role in my professional development, and the unique perspective you bring to the table has always served to improve any project we have ever collaborated on. Thank you for your diligence, patience, and encouragement.

Daniel Ruzzante - Thank you for your contributions as a member of my graduate committee to this research, and for your patience; particularly over the past several months as I endeavored to complete this thesis in record time. The reality checks you've provided have helped to keep me grounded, and made me recognize the folly of setting goals that are overambitious; valuable lessons for the future.

The field work required for this research relied on a host of volunteers from governmental agencies in both the United States and Canada, as well as private citizens too many to name here individually, so I thank them all collectively. However additional field assistance was provided by M. Coulson, D. Hardie, D. Fraser, J. Seibert, V. Yaroshewsky, C. Purchase, S. Clifford, L. Weir, and C. Minto. Laboratory assistance was provided by I. Paterson, L. Anstey, and L. Baillie.

Finally, deepest thanks to my wife Jenny. Despite this degree taking longer than anticipated, you've always been there with love and encouragement. Thanks for your unwavering support and patience through the most difficult of times.

CHAPTER 1

INTRODUCTION

1.1 SPATIAL DISTRIBUTION OF NEUTRAL GENETIC VARIATION

The long-term persistence of species is to a great extent dependent on the capacity of populations to mount variable responses to unpredictable perturbations. The breadth of this response is largely contingent on the amount of heritable variation present at the population level (Allendorf and Lundquist 2003). Although maintenance of genetic variation and connectivity among populations from across the range of habitats in which a species persists can increase the capacity for evolutionary change (Mace and Purvis 2008), populations are not necessarily equivalent in their levels of genetic variation, or in their responses to future environmental conditions (Petit et al. 1998). Thus, resolving the magnitude and spatial distribution of genetic variation from across a species' range is an integral component of resource management, and is important to the identification and prioritization of populations for focusing conservation resources, the designation of management units, the protection of adaptive genetic variation, and the preservation of species adaptive potential (Allendorf 1986). Neutral molecular markers, such as mitochondrial DNA (mtDNA) and microsatellites, figure prominently in this regard, because estimates of genetic variation revealed through the examination of neutral markers are often used as proxies of genome wide variation (but see Väli et al. 2008; Ljungqvist et al. 2010). Although the single, non-recombining mtDNA molecule cannot generate independently replicated data (Ballard and Whitlock 2004) for examination of

spatial genetic patterns, multiple unlinked microsatellite loci can provide independent data for assessing the magnitude and spatial distribution of genetic variation, and are increasingly sought to inform conservation and management decisions (Crandall et al. 2000; Allendorf et al. 2004).

The magnitude and spatial distribution of neutral genetic variation for most species reflects a complex mosaic of historical demographic factors, microevolutionary processes, and anthropogenic influences; the relative effects of each of which are difficult to discern. The spatial distribution of neutral genetic variation for most species is heavily influenced by historical demographic factors associated with climatic change (Beaumont 1999; Storz and Beaumont 2002; Bos et al. 2008). Recurrent glacial cycles during the Pleistocene epoch likely significantly reduced intraspecific levels of genetic variation for many temperate species through a combination of successive population bottleneck and founder events (Hewitt 2000); different modes of post-glacial dispersal (Ibrahim et al. 1996) during interglacial range expansions influenced the initial spatial structure of populations. Irrespective of the significant role of historical demographic factors, knowledge of the effects of microevolutionary processes and anthropogenic influences is essential to the understanding of spatial genetic variation (Bohonak 1999; Gagnon and Angers 2006 and references therein). Contemporary patterns of genetic variation may reflect the processes of gene flow and/or genetic drift on the distribution of gene frequencies since post-glacial colonization, but may be distorted by habitat fragmentation or translocations of individuals among demes.

1.2 AMERICAN SHAD AS A MODEL SPECIES

The American shad (*Alosa sapidissima* Wilson, 1811) (hereafter shad) is an anadromous clupeid indigenous to the Atlantic coast of North America. Shad exhibit a broad native distribution (~30°N-50°N), with spawning runs distributed among major rivers from Florida and Quebec (Walburg and Nichols 1967). Across this range, shad exhibit intraspecific life history variation, and a dichotomous reproductive strategy. Spawning runs south of Cape Hatteras, NC, are entirely semelparous (spawn once and die), while those to the north exhibit an increasing iteroparity with latitude (Leggett and Carscadden 1978). Juveniles emigrate from natal rivers and shoal as mixed assemblages along coastal regions until sexual maturity (Talbot and Sykes 1958; Dadswell et al. 1987). Available evidence indicates that reproductive isolation among spawning runs is maintained through a combination of philopatry (Hendricks et al. 2002) and spawning site fidelity (Melvin et al. 1986). During the last glacial maximum (LGM; 23-18 ky bp), shad are thought to have been confined to the southeastern United States (Bentzen et al. 1989). Post-glacial dispersal from this region may have followed a stepping stone process (Ibrahim et al. 1996) as habitat may have become sequentially suitable for colonization, and may have manifested as successive reductions in genetic variation with increasing latitude.

Previous studies of shad neutral genetic variation have revealed that individual rivers support genetically distinguishable spawning runs (Bentzen et al. 1989), and provide powerful management tools for resolving the spatial distribution of genetic variation, and scale of population structure. Prior synoptic surveys of shad genetic variation used mtDNA to resolve population structure (Bentzen et al. 1989; Epifanio et

al. 1995; Waldman et al. 1996). Although highly polymorphic microsatellite loci generally provide a greater degree of resolution for this purpose, previous microsatellite based studies of shad have been limited to few loci and few populations (Brown et al. 2000; Waters et al. 2000). Further, previous studies of shad genetic variation have tended to focus on spawning runs in the United States; only five of the more than 20 suspected spawning runs in Canada (Chaput and Bradford 2003 and references therein) have ever been assessed.

Shad are of increasing conservation concern, as range-wide declines in population abundance attributable to anthropogenic factors (i.e. overfishing, dams, pollution; Bilkovic et al. 2002a), and substantive reductions in the number of extant spawning runs (Limburg et al. 2003), have resulted in moratoria being imposed upon once profitable fisheries (ASMFC 1999 as in Bilkovic et al. 2002b). Substantial resources have been allocated to the restoration of U.S. shad spawning runs (e.g. supportive breeding, out of basin stock transfers, modification of fish passage, dam removal) with varying success (reviewed in Hendricks 2003; Cooke and Leach 2003; Olney et al. 2003; St. Pierre 2003; Weaver et al. 2003). Understanding the spatial distribution of shad genetic variation may benefit further conservation/restoration efforts, as knowledge of the relative contribution of populations to species level genetic variation could aid the prioritization of populations for management, leading to an effective conservation strategy, and the long-term persistence of shad.

1.3 THESIS OVERVIEW

The context for this thesis is applied conservation genetics. The objective of this thesis is to examine how knowledge of the magnitude and spatial distribution of neutral

genetic variation from across a species range can benefit management goals, conservation objectives, and the understanding of a species' evolutionary history.

Chapter 2 explores known spatial patterns of neutral genetic variation within and among 49 Nearctic freshwater and anadromous fishes to assess whether estimates of genetic diversity obtained from only a portion of a species' range are valid for inferences of broad spatial scale patterns both within and among species. Results from this chapter suggest that latitude is an important determinant of the amount and spatial distribution of genetic variation for many Nearctic fishes. This study also suggests that estimates of genetic variation obtained from a portion of a species' range limits on inference of broader (i.e. rangewide) spatial patterns, and that rangewide data is in fact required for adequate representation and interpretation of broad spatial patterns of neutral genetic variation.

The remainder of this thesis focuses on the magnitude and spatial distribution of neutral genetic variation among shad spawning runs from across the species native distribution. Chapter 3 uses 13 microsatellite loci to examine the partitioning of genetic variation within four and among 12 Canadian spawning runs in order to resolve the spatial scale of population structure within that portion of the species' range, and to identify the number of genetically distinguishable populations (i.e. evolutionary paradigm; Waples and Gaggiotti 2006). Results suggest evidence for shad metapopulation structure, and suggest that fisheries managers should be concerned with the loss of shad genetic variation on both river and regional scales.

Chapter 4 examines the spatial partitioning of genetic variation among 33 populations from across the species range. Although some analyses used in chapter 3 are

repeated here, the goal of this chapter is not to resolve the spatial scale of population structure, but to determine whether different patterns of population structure are evident across the species range, and whether these patterns are consistent with the influence of historical demographic factors, contemporary microevolutionary processes, or anthropogenic effects. This chapter further assess i) whether the distribution of genetic variation across formerly glaciated habitats is consistent with stepwise post-glacial colonization, and ii) the relative contribution of each population to overall species genetic variation. Results indicate that different patterns of population structure are evident across the species range that may reflect the influences of historical population bottlenecks/founder events, gene flow, and stocking practices on the spatial distribution of genetic variation. Results from this chapter also indicate that post-glacial colonization likely progressed as a stepwise process, and that U.S. and Canadian spawning runs exhibit reciprocal patterns of genetic variation and differentiation.

The final chapter (chapter 5) synthesizes the major findings of the research conducted in this thesis, discusses the implications of these results for the conservation of shad genetic variation, and provides an outlook for future shad research.

1.4 PUBLICATIONS ARISING FROM THE THESIS

At this time one publication has been generated from this thesis (chapter 3):

Hasselman, D.J., R.G. Bradford, and P. Bentzen. 2010. Taking stock: defining populations of American shad (*Alosa sapidissima*) in Canada using neutral genetic markers. Canadian Journal of Fisheries and Aquatic Sciences 67: 1021-1039.

The publisher of Canadian Journal of Fisheries and Aquatic Sciences (NRC Press) does not require copyright permission for authors to reproduce their studies as parts of theses.

CHAPTER 2

SPATIAL DISTRIBUTION OF NEUTRAL GENETIC VARIATION IN NEARCTIC FISHES: DOES LATITUDE MATTER?

2.1 INTRODUCTION

Knowledge of the spatial distribution of neutral genetic variation can bolster understanding of species evolutionary histories, and facilitate conservation planning. While information about the relative influences of anthropogenic factors and microevolutionary processes contribute to our understanding of contemporary patterns of genetic variation (Bohonak 1999; Gagnon and Angers 2006 and references therein), the spatial distribution of genetic diversity for most species is heavily influenced by historical demographic factors associated with climatic change (Beaumont 1999; Storz and Beaumont 2002; Bos et al. 2008). Recurrent glacial cycles during the Pleistocene epoch significantly affected the spatial distributions of many extant lineages. Species physically displaced by advancing glaciers may have dispersed to new locations, or survived in refugia near glacial margins and recolonized habitat that became available during interglacial periods (~10,000-12,000 yr duration) (Dawson 1992; Hewitt 2000). The net effect of the presumed resultant population bottlenecks/founding events associated with Pleistocene glaciations ultimately diminished intraspecific levels of genetic variation.

In North America, the most recent major Pleistocene glaciation, the Wisconsinan (~30,000-10,000 ybp), had a pronounced effect on the distributions of species. During the last glacial maximum (LGM, 23,000-18,000 ybp) the Laurentide ice sheet exceeded the combined size of glaciers located in Europe and Asia (Dawson 1992), and the glacial

margin extended south along the Atlantic coast to Long Island (~40°N; Schmidt 1986). Aquatic species were particularly affected, as dispersal routes away from the advancing glacier were limited to direct connections formed among freshwater habitats. Loss of genetic variation presumably accompanied the destruction or displacement of many northern aquatic populations. Further losses may have resulted from biome compression at glacial margins (i.e. reductions in habitat size and availability due to species' restriction in glacial refugia; Pielou 1991), and the depressed evolutionary effective population size experienced over millennia (Bernatchez and Wilson 1998). Following the onset of glacial recession ~15,000 ybp (Dyke and Prest 1987; Hewitt 2000), the creation of large pro-glacial lakes and the freshening of coastal ocean waters by rapidly melting glaciers facilitated post-glacial dispersal and colonization of newly available habitat over vast geographical ranges, further reducing intraspecific genetic diversity.

Among aquatic taxa, the impacts of Pleistocene glaciations on the spatial distribution of genetic variation has been explored most thoroughly for freshwater fishes (e.g. Bernatchez and Dodson 1991; Wilson and Hebert 1998; Turgeon and Bernatchez 2001). Anadromous fishes were likely similarly affected by Pleistocene glaciations because of their reliance on freshwater habitat for reproduction. However, their dispersal into marine waters between reproductive events suggests that the spatial distribution of genetic diversity may be interpreted in the absence of confounding factors such as those arising from isostatic rebound (but see Bernatchez and Dodson 1990, Bernatchez 1997) that can alter drainage patterns, and bring into secondary contact populations that experienced different demographic histories in separate refugia (e.g. Bernatchez and Dodson 1991; Wilson and Hebert 1998).

In their seminal paper, Bernatchez and Wilson (1998) reviewed the distribution of genetic variation for freshwater and anadromous Palearctic and Nearctic fishes using mtDNA. For each of 42 species, the authors provided a global estimate of intraspecific genetic diversity in the form of a relative index of nucleotide diversity, and related this value to the median latitude for the contemporary distribution of each species. Their study revealed a pattern of decreasing genetic diversity with latitude, and using piecewise linear regression revealed a break at 46°N; generally consistent with the median latitude of the southern limit of Pleistocene glaciations (44°N; Fulton and Andrews 1987) (see Figure 3 in Bernatchez and Wilson 1998). A re-assessment by McDowall (1999) revealed that diadromous fishes generally possessed low levels of nucleotide diversity, and suggested that there was no latitudinal effect on genetic variation among diadromous species (see Figure 1 in McDowall 1999). In response, Bernatchez and Wilson (1999) used nonlinear piecewise regression to support a latitudinal decline of genetic variation (r^2=0.67) among diadromous fishes. However, a fundamental premise of the original analysis conducted by Bernatchez and Wilson (1998), that species median latitude is a suitable measure of species distribution, may not have been appropriate for inference of broad spatial patterns of genetic variation given the nature of the available data. Estimates of global nucleotide diversity for most of the species considered were obtained from studies that examined only a portion of the species range, in some cases the northern or southern extremes of their distribution (e.g. Atlantic sturgeon *Acipenser oxyrhynchus*, Bowen and Avise 1990). Post-glacial colonization for some of these species may have followed a stepwise process (Ibrahim et al. 1996), with successive population founder events leading to sequential reductions in genetic diversity with increasing latitude.

Thus, the latitudinal distribution of genetic variation within species may mirror the spatial patterns observed among species. As a consequence, condensing genetic data from studies that sampled only a portion of species latitudinal ranges to a single datum may lead to potential inaccuracies in among species comparisons. In particular, this approach may bias global intraspecific diversity estimates either upward or downward, depending on whether data were derived primarily from the southern or northern portion of a species range. Greater accuracy and more meaningful interpretation of broad spatial patterns of genetic variation may be achieved through the examination of the mean latitude of the populations from which estimates of intraspecific diversity are available.

There are additional factors in Bernatchez and Wilson (1998) that may have further confounded patterns of spatial genetic variation. Their inclusion of hatchery strains (i.e. rainbow trout *Oncorhynchus mykiss*, Danzmann et al. 1993; masu salmon *Oncorhynchus masu*, Kijima and Matsunami 1992; Atlantic salmon *Salmo salar*, Bermingham et al. 1991) does not permit accurate geographic representation for estimates of genetic variation relative to natural populations, and is not appropriate for comparisons among naturally occurring populations of other species. Their failure to consider separately populations of species which constitute different mtDNA lineages (may have experienced separate demographic histories in isolated glacial refugia) (i.e. lake whitefish *Coregonus clupeaformis*, Bernatchez and Dodson 1991; lake trout *Salvelinus namaycush*, Wilson and Hebert 1996) does not take into account the potential effects of secondary contact and introgression following post-glacial colonization on estimates of genetic variation for some populations. Their inclusion of species characterized for different regions of the mitochondrial genome (i.e. threespine

stickleback *Gasterosteus aculeatus*, Orti et al. 1994; northern redbelly dace *Phoxinus eos* Toline and Baker 1995; brown trout *Salmo trutta* Bernatchez et al. 1992, Giuffra et al. 1994; sockeye salmon *Oncorhynchus nerka*, Bickham et al. 1995; chum salmon *Oncorhynchus keta*, Cronin et al 1993; chinook salmon *Oncorhynchus tshawytscha*, Cronin et al. 1993) does not account for the effect that different substitution rates may have on estimates of genetic variation, and the subsequent influence in among species comparisons.

Additional genetic data has been generated for many of the species considered by Bernatchez and Wilson (1998) since the publication of their work, which provides the opportunity to revisit spatial patterns of genetic variation for freshwater and anadromous fishes. To this end I first examined whether species median latitude was representative of the mean distribution of the populations from which estimates of global genetic diversity were originally obtained. Because of the potential for different glacial histories of Europe and North America (and associated alternate routes of post-glacial colonization; Hewitt 1996, 2000) to confound latitudinal patterns of genetic variation, I restricted subsequent analyses to Nearctic fishes, but expanded the dataset available for this group by conducting a literature survey for each Nearctic species examined by Bernatchez and Wilson (1998) to include additional populations (thereby increasing geographic coverage), and to incorporate data from nuclear markers (allozymes and microsatellites). Analysis of nuclear genetic variation may provide a greater understanding of the influence of historical demographic events on spatial patterns of genetic variation, as multiple unlinked nuclear markers can provide independent data about population histories, and can retain high levels of polymorphism in founder populations (Bonhomme

et al. 2008). For this portion of the analysis I also included species not considered by Bernatchez and Wilson (1998) to increase the taxonomic coverage of the study.

2.2 Materials and Methods

2.2.1 Assessment of Spatial Genetic Variation

Evaluation of species median latitude

With the exceptions of rainbow trout, masu salmon, and Atlantic salmon (hatchery strains), I examined the source studies for each species (N=39) listed in Appendix I of Bernatchez and Wilson (1998), and estimated the mean distribution (±95% CI; SYSTAT v.11, SPSS 2000) of the populations originally assessed to understand the geographic range across which estimates of global intraspecific genetic diversity were applicable. I then determined whether the median latitude reported for each species fell outside of this range.

Spatial genetic variation among Nearctic fishes

Where possible, I estimated population level haplotypic (nucleon) diversity (h) from the original source studies following Nei and Tajima (1981) to evaluate the diversity of mtDNA lineages within populations, and the spatial distribution of intraspecific genetic variation for Nearctic fishes. Haplotypic diversity was calculated as:

$$h = n\left(1 - \sum_{i=1}^{r} x_i^2\right)/(n-1),$$

where x_i is the frequency of the ith type of mtDNA in a population of n specimens, and r is the number of mtDNA types. Although this index is usually applied to the

heterozygosity at nuclear loci (for randomly mating diploid organisms), it is an appropriate measure for the diversity of maternal lineages (Nei 1987). However, comparisons among different studies using this metric are difficult, as haplotypic diversity is influenced by the types and numbers of restriction enzymes analyzed, region of the mitochondrial genome surveyed, total number of base pairs examined, and analytical method employed (i.e. RFLP (restriction fragment length polymorphism) or SSCP (single stranded conformation polymorphism)) (Nei and Tajima 1981). Therefore, I examined spatial patterns of genetic variation among 36 Nearctic fishes using the relative index of nucleotide diversity reported by Bernatchez and Wilson (1998) and species mean distributions estimated from the original source studies. This index was estimated by multiplying the deepest branch length of intraspecific trees by each species overall haplotypic diversity. Bernatchez and Wilson (1998) found that this index approximated nucleotide diversity (π; Nei and Tajima 1981), which considers the frequency of occurrence of mtDNA types in a population as well as the divergence among mtDNA lineages (Brown et al. 1992), and can be considered a measure of heterozygosity at the nucleotide level in randomly mating populations (Nei 1987). I evaluated associations between latitude and genetic diversity indices using Pearson correlation coefficients, and regression analyses (SYSTAT v.11; SPSS 2000).

2.2.2 Literature Survey

I conducted a search of the primary literature current to April 2009 using the Web of Science® database (ISI Web of knowledge®) for titles, abstracts, and keywords for all possible combinations of species common names and Latin binomials with: phylogeography, biogeography, zoogeography, latitude, glaciation, dispersal capabilities,

genetic diversity, microsatellite(s), anadromous, freshwater, and marine. I expanded the search beyond the Nearctic fishes considered by Bernatchez and Wilson (1998) to include 11 additional species (three freshwater, eight anadromous). In total, the search criteria encompassed 49 Nearctic species, and yielded 329 articles; 125 of which contained data suitable for exploring spatial patterns of genetic variation (Appendix 1).

Spatial distribution of intraspecific genetic variation

I examined intraspecific patterns of genetic variation from mtDNA, allozyme, and microsatellite based studies separately for each species, as interpretations of genetic diversity from these markers may depend on the number of base pairs/loci surveyed, sample sizes, and the region of the mtDNA genome/loci assessed (Ferguson and Danzmann 1998). Because different portions of the mitochondrial genome are known to evolve at different rates (Marshall et al. 2009), I considered separately those studies which examined the control region and those which examined remaining portions of the mtDNA genome (e.g. cytochrome b, NADH). I also considered separately those studies of mtDNA which employed RFLP and SSCP based analytical methods. Theoretically, diversity estimates from RFLP and sequence based studies should not differ (Ferguson and Danzmann 1998). However, if restriction enzymes were chosen *a priori* because of the level of polymorphism detected, RFLP analyses could generate higher diversity estimates than sequence based studies, which do not omit portions of sequence that are not polymorphic.

Where possible I collected data on nucleotide diversity, and estimated population level haplotypic diversity from mtDNA based studies following Nei and Tajima (1981). I

collected observed heterozygosity (H_O) for allozyme studies, and H_O and allelic richness (R) for microsatellite based studies. Because sample size can influence the amount of diversity detected, I restricted analyses to populations which had been genotyped across at least 10 individuals for mtDNA based studies, and at least 30 individuals for nuclear genome based studies (allozymes, microsatellites). Although a sample size of 50-100 individuals has been advocated for the analysis of genetic distance and population structure using microsatellites (Ruzzante 1998), application of this criterion eliminated >65% of the populations assessed across all studies. A minimum sample size of 30 individuals retained sufficient numbers of populations to assess spatial patterns of intraspecific variation, while removing data based on small sample sizes; resampling of microsatellite data for 30 American shad (Hasselman unpubl. data) revealed similar estimates of diversity (H_O, R) (data not shown). Where appropriate, I combined data across similar studies (mtDNA region, analytical method) using weighted arithmetic means (similar to Taylor and Bentzen 1993) to account for differences in sample sizes among studies assessing the same populations (e.g. Atlantic sturgeon, Wirgin et al. 2000, 2007). I considered separately those populations that belong to different mtDNA lineages and which may have experienced separate demographic histories in isolated glacial refugia (e.g. sympatric pairs with allopatric origins). Where secondary contact and introgression among phylogenetic lineages has occurred, I classified populations to the lineage which comprised the greatest proportion of the genomic composition. I evaluated associations between latitude and genetic diversity indices using Pearson correlation coefficients, and regression analyses (SYSTAT v.11; SPSS 2000).

Spatial distribution of genetic variation among species

To determine whether broad spatial patterns of genetic variation observed within species were reflected among species, I examined the spatial distribution for estimates of global nucleotide diversity and nuclear genetic variation among species using the mean latitude of the populations from which these estimates were obtained. Pearson correlation coefficients and regression analyses (SYSTAT v.11; SPSS 2000) were used to evaluate associations between latitude and genetic diversity indices.

2.3 RESULTS

2.3.1 ASSESSMENT OF SPATIAL GENETIC VARIATION

Evaluation of species median latitude

Species median latitudes reported by Bernatchez and Wilson (1998) fell outside of the mean (±95% CI) distribution of the populations considered in source studies for 20 of the 39 fishes examined (Figure 2.1). The use of mean study latitudes shifted estimates of global nucleotide diversity for 15 of these 20 species to lower latitudes by 4.21° on average. In some instances species median latitudes fell just outside the range of the populations originally assessed (i.e. <1° difference) (e.g. redear sunfish, white sucker, lake whitefish). However, there was a ~9° difference between species median latitude and the range of Atlantic sturgeon and lake trout populations examined by Bowen and Avise (1990) and Wilson and Hebert (1996), respectively (Figure 2.1).

The source studies listed in Appendix I of Bernatchez and Wilson (1998) allowed estimation of haplotypic diversity among populations for 22 species of Nearctic fishes. Negative correlations of intraspecific genetic variation with latitude were observed for 17

species (mean r=-0.40±0.28; Table 2.1), whereas positive correlations were observed for four species: chum salmon, chinook salmon, cisco, and white sucker. Two studies of lake sturgeon were cited in Bernatchez and Wilson (1998), one of which exhibited a negative correlation of intraspecific genetic variation with latitude, and the other a positive correlation. Cumulatively, these results support the notion that latitude is an important determinant of the spatial distribution of intraspecific genetic variation for many Nearctic fishes, and that the prior use of species median latitude may not be representative of the populations from which these estimates were originally obtained. Although the use of different measures may have a small effect on the latitude considered to be 'representative' of global genetic diversity for any given species, the cumulative effects when accrued across species may have implications for our understanding of broad spatial patterns of genetic variation. These results suggest that that a reassessment of the spatial distribution of genetic variation among Nearctic fishes is warranted.

Spatial genetic variation among Nearctic fishes

Using mean latitude of source study populations for 36 Nearctic fishes, negative exponential regression (Figure 2.2a) and linear regression models were highly significant ($p<0.001$), and explained a substantial proportion of variation (r^2=0.38 and 0.28, respectively) in the distribution of nucleotide diversity, with greater variation observed among southern latitude species than northern latitude species (Figure 2.2b). This result contrasts with that of Bernatchez and Wilson (1998), who found that 87% of variation among 42 Nearctic and Palearctic species was explained by a piecewise linear regression incorporating a break at 46°N when using species median latitudes. However, I could not

Figure 2.1 Median latitude (★) for 39 Nearctic and Palearctic fishes reported in Appendix I of Bernatchez and Wilson (1998) versus mean latitude (●) of the populations examined in the original source studies (distribution bars: ±95% CI). Number of populations assessed in the original source studies for each species are displayed.

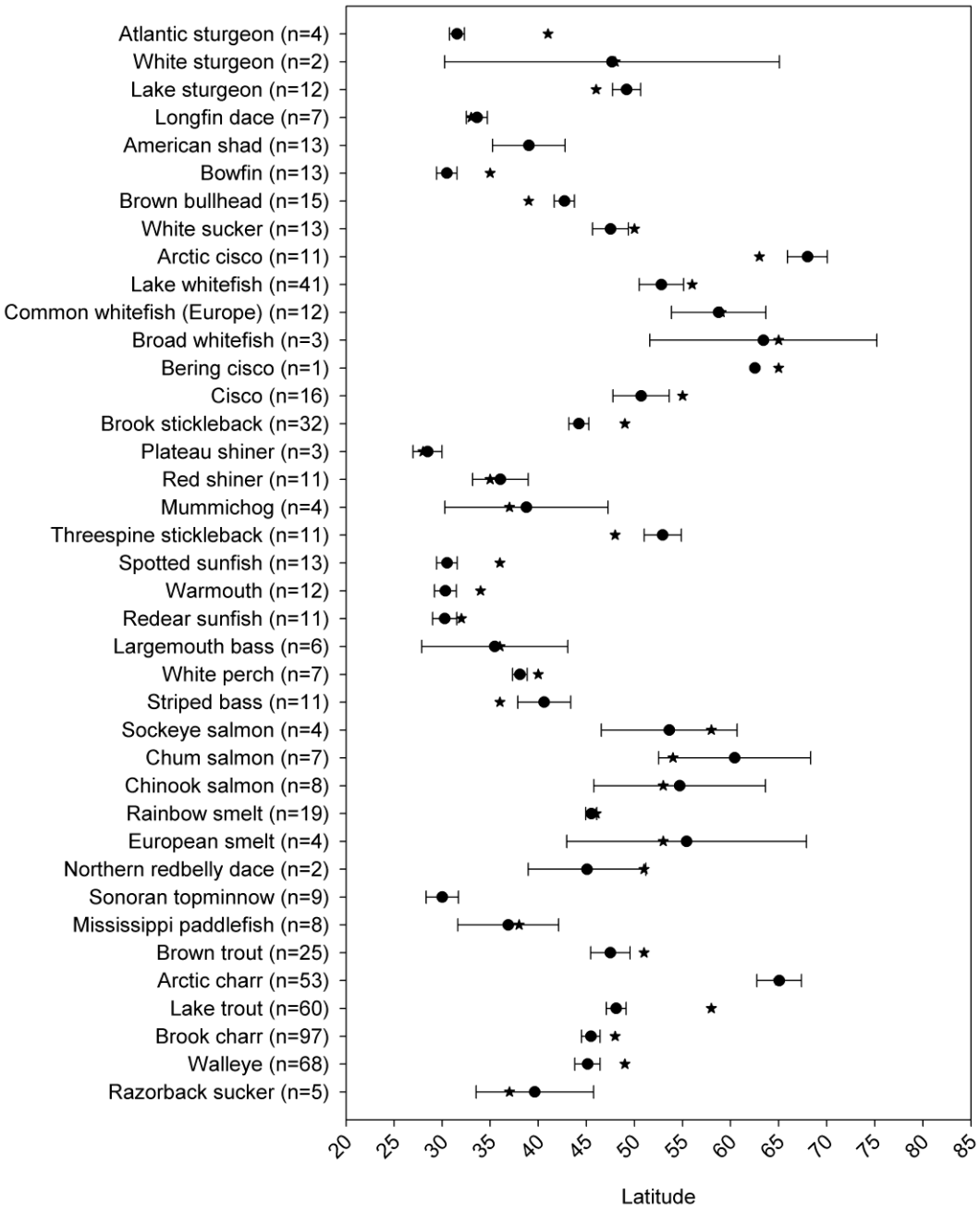

Atlantic sturgeon (n=4)
White sturgeon (n=2)
Lake sturgeon (n=12)
Longfin dace (n=7)
American shad (n=13)
Bowfin (n=13)
Brown bullhead (n=15)
White sucker (n=13)
Arctic cisco (n=11)
Lake whitefish (n=41)
Common whitefish (Europe) (n=12)
Broad whitefish (n=3)
Bering cisco (n=1)
Cisco (n=16)
Brook stickleback (n=32)
Plateau shiner (n=3)
Red shiner (n=11)
Mummichog (n=4)
Threespine stickleback (n=11)
Spotted sunfish (n=13)
Warmouth (n=12)
Redear sunfish (n=11)
Largemouth bass (n=6)
White perch (n=7)
Striped bass (n=11)
Sockeye salmon (n=4)
Chum salmon (n=7)
Chinook salmon (n=8)
Rainbow smelt (n=19)
European smelt (n=4)
Northern redbelly dace (n=2)
Sonoran topminnow (n=9)
Mississippi paddlefish (n=8)
Brown trout (n=25)
Arctic charr (n=53)
Lake trout (n=60)
Brook charr (n=97)
Walleye (n=68)
Razorback sucker (n=5)
20
25
30
35
40
45
50
55
60
65
70
75
80
85
Latitude

Table 2.1 Correlation of haplotype diversity (*h*) with latitude among populations within 22 species of Nearctic fishes using data from source studies listed in Appendix I of Bernatchez and Wilson (1998).

Species	Source study	Populations (N)	Pearson correlation
Lake sturgeon	Ferguson et al. 1993	8	-0.482
Lake sturgeon	Guenette et al. 1993	4	0.918
Longfin dace	Tibbets and Dowling 1996	7	-0.690
American shad	Bentzen et al. 1989	13	-0.257
Brown bullhead	Murdoch and Hebert 1997	15	-0.175
White sucker	Lafontaine and Dodson 1997	13	0.519
Arctic cisco	Bickham et al. 1992	11	-0.173
Cicso	Bernatchez and Dodson 1990	8	0.276
Cicso	Snyder et al. 1992	8	0.230
Red shiner	Richardson and Gold 1995a	11	-0.522
Mummichog	Gonzales-Villasenor and Powers 1990	4	-0.667
Threespine stickleback	O'Reilly et al. 1993	10	-0.072
Largemouth bass	Nedbal and Phillip 1994	6	-0.336
White perch	Mulligan and Chapman 1989	7	-0.744
Striped bass	Wirgin et al. 1993b	5	-0.655
Sockeye	Bickham et al. 1995	4	-0.630

Table 2.1 *continued*

Species	Source study	Populations (N)	Pearson correlation
Chum	Cronin et al. 1993	7	0.541
Chinook	Cronin et al. 1993	8	0.616
Rainbow smelt	Taylor and Bentzen 1993	19	-0.046
Sonoran topminnow	Quattro et al. 1996	9	-0.881
Mississippi paddlefish	Epifanio et al. 1996	8	-0.04
Lake trout	Wilson and Hebert 1996	60	-0.017
Brook charr	Danzmann et al. 1998	97	-0.264
Razorback sucker	Dowling et al. 1996	5	-0.6

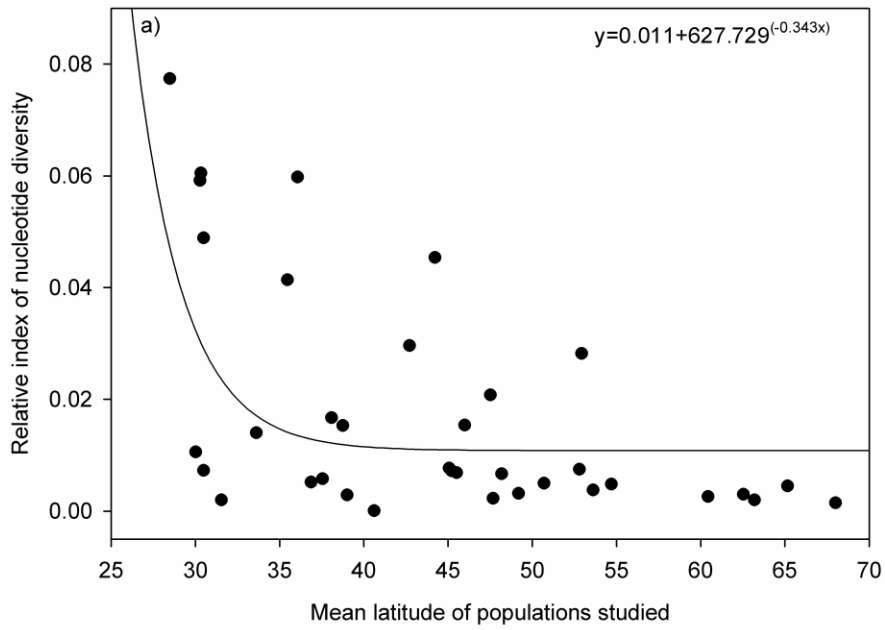
replicate this result using their full data set (i.e. 42 species, nucleotide diversity, species median latitude), as only 35.2% of the variation in the distribution of nucleotide diversity was explained with a break at 46°N. Piecewise linear regression on their full data set instead revealed that the greatest proportion of variation (41.9%) was explained by a break at 38°N, more consistent with the LGM (data not shown). Although the source of this discrepancy is uncertain, visual inspection of their Figure 3 suggests that the proportion of variation explained with piecewise linear regression may be lower than 87%. Regardless, these results suggest that the relationship between mean study latitude and nucleotide diversity among the 36 Nearctic fishes examined herein is best characterized by a negative exponential function (Figure 2.2a).

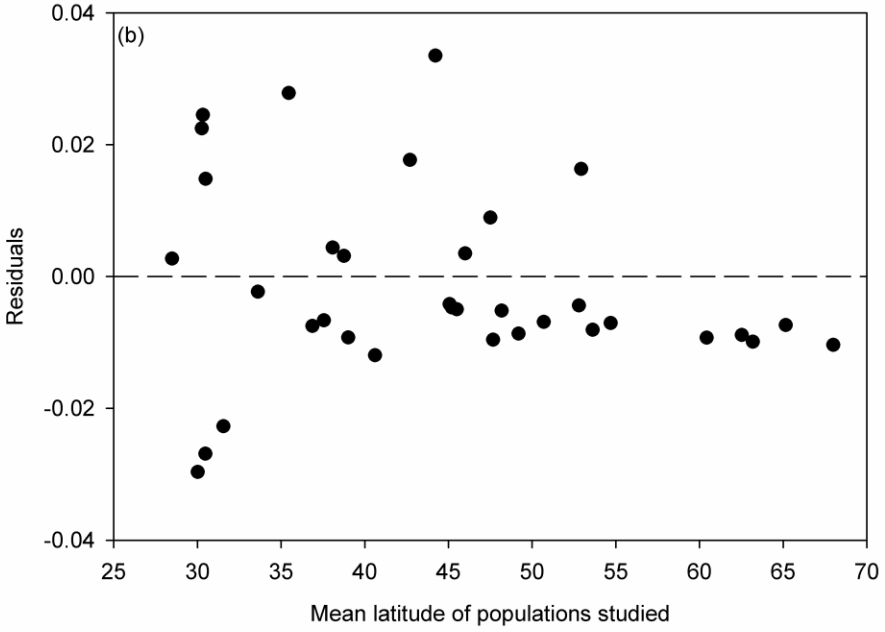
Freshwater vs. anadromous species

Following McDowall (1999), freshwater and anadromous species were differentiated to further explore patterns of genetic variation among Nearctic fishes. In his modification of Bernatchez and Wilson (1998)'s Figure 3, McDowall (1999) categorized mummichogs as a freshwater species although the original source study (Gonzales-Villasenor and Powers 1990) examined estuarine populations. Similarly, brook charr, threespine stickleback, and lake whitefish were considered diadromous, although the original source studies examined freshwater populations (Danzmann et al. 1998, O'Reilly et al. 1993, Bernatchez and Dodson 1991). These species designations were re-assigned freshwater status to more accurately reflect the life histories of the populations considered in the source studies prior to analyses.

Declines in nucleotide diversity with latitude were greater among freshwater

Figure 2.2 Spatial patterns of genetic variation among 36 Nearctic fishes using mean study latitude of the populations assessed from original source studies: a) a negative exponential function characterized the latitudinal decline of nucleotide diversity among species, b) residuals from negative exponential regression showed decreasing variation with increasing latitude.


a)
y=0.011+627.729(-0.343x)
Relative index of nucleotide diversity
0.08
0.06
0.04
0.02
0.00
25
30
35
40
45
50
55
60
65
70
Mean latitude of populations studied


(b)
Residuals
0.04
0.02
0.00
-0.02
-0.04
25
30
35
40
45
50
55
60
65
70
Mean latitude of populations studied

species than anadromous fishes (Figure 2.3), a result also observed by McDowall (1999). Negative exponential regression (Figure 2.3a) and linear regression models explained a significant ($p<0.01$) and substantial proportion of variation (0.30 and 0.28, respectively) in the distribution of genetic diversity among freshwater species. However, this negative exponential trend was sensitive to the inclusion of desert species and non-teleost fishes. Desert fishes of the southwestern U.S. generally constituted those low latitude species that exhibited low levels of global nucleotide diversity. When two endangered desert species (Sonoran topminnow, razorback sucker) were removed from analysis, linear regression was significant ($p=0.002$) and explained a greater proportion of variation ($r^2=0.40$) than the negative exponential function (Figure 2.3a). Bernatchez and Wilson (1998) noted that chondrostean and holostean fishes exhibited low levels of diversity, and suggested that this may be due to slower mtDNA mutation rates in non-teleosts, and/or differences in evolutionary female effective population size between teleost and non-teleost fishes (Krieger and Fuerst 2002, Avise et al. 1992). When lake sturgeon, bowfin and Mississippi paddlefish were removed from analysis, linear regression ($y=-0.002x+0.121$) was highly significant ($p<0.001$), and explained a greater proportion of variation ($r^2=0.61$) than the negative exponential function, or the exclusion of desert fishes alone (Figure 2.3a).

Anadromous fishes exhibited less variation for global estimates of nucleotide diversity with latitude than freshwater species (Figure 2.3b). Linear regression analysis revealed a non-significant ($r^2=0.17$; $p>0.05$) effect of latitude on the spatial distribution of nucleotide diversity among anadromous fishes (Figure 2.4b). This result is consistent with that obtained by McDowall (1999), and as noted by Bernatchez and Wilson (1999)

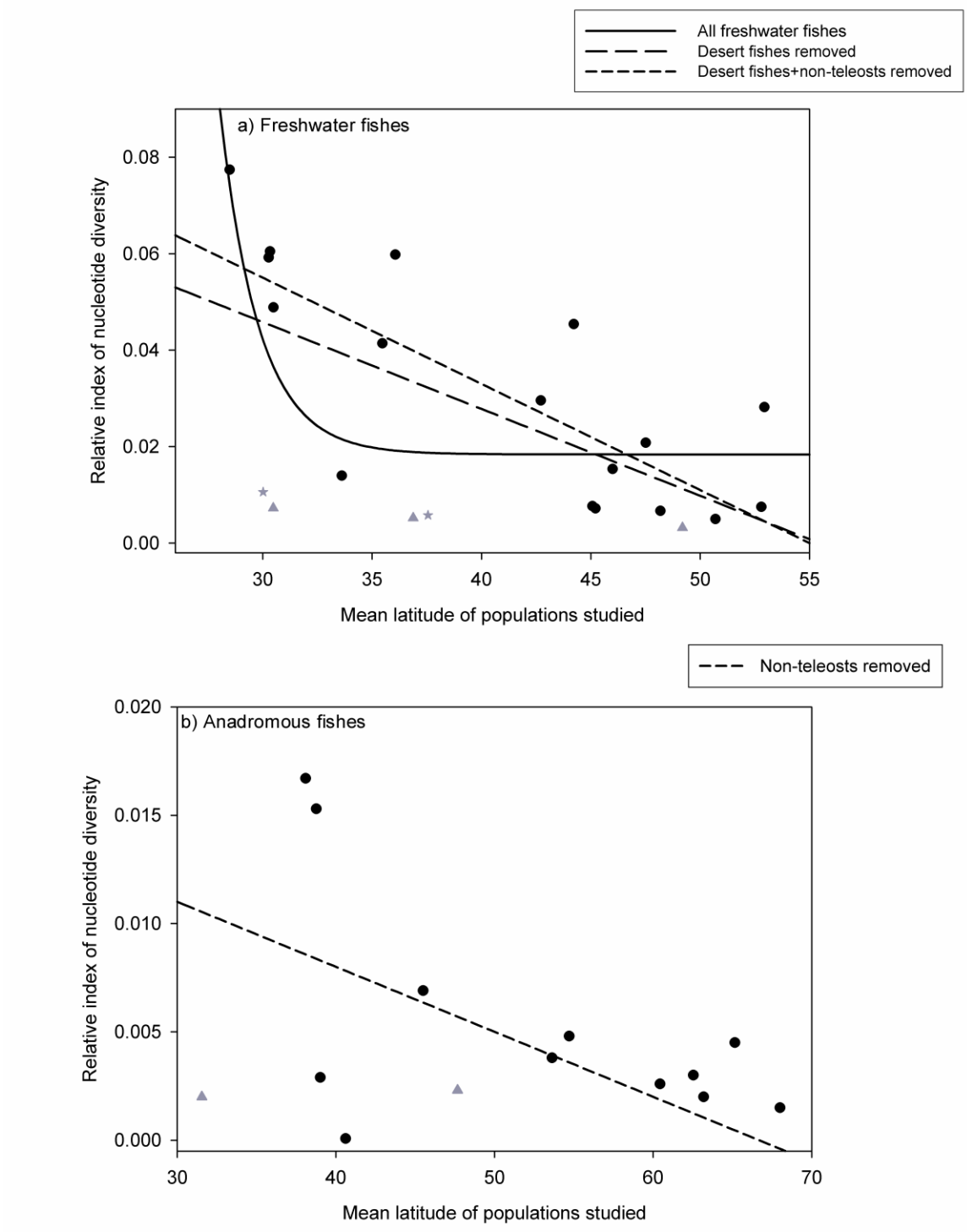
may have been due to the low nucleotide diversity observed among anadromous species found at low latitudes. Using non-linear piecewise regression, Bernatchez and Wilson (1999) observed that latitude explained 67% of the variance in nucleotide diversity, and detected a significant latitudinal effect when sturgeon were excluded ($p=0.02$). Similarly, omission of sturgeons from analysis in this study revealed a significant ($r^2=0.35$; $p<0.05$) latitudinal effect (i.e. linear decline; $y=-0.0003x+0.02$) among remaining anadromous species (Figure 2.3b).

2.3.2 LITERATURE SURVEY

Spatial distribution of intraspecific genetic variation

Of the 49 Nearctic fishes surveyed, spatial patterns of intraspecific genetic variation could be examined for 42 species. Due to sample size limitations and lack of geographic coverage, latitudinal trends of intraspecific genetic diversity could not be assessed for white sturgeon, bowfin, warmouth, redear sunfish, brook stickleback, Bering cisco, or broad whitefish. Of those Nearctic species that could be assessed, extensive variation in correlations of genetic diversity with latitude were observed among diversity indices, marker types, and analytical methodologies (Appendix 2). Some of this variation may reflect the effects of different molecular markers, methodologies or populations considered. For example, Atlantic sturgeon exhibited negative correlations with latitude for microsatellite based heterozygosity (-0.72) and allelic richness (-0.15), as well as haplotype diversity estimated from mtDNA D-loop sequences (-0.86), but a positive correlation with latitude for nucleotide diversity estimated from D-loop sequences (0.98) (Appendix 2). However, this estimate of nucleotide diversity was obtained from only

Figure 2.3 Spatial patterns of genetic variation among a) freshwater species (n=22), and b) anadromous species (n=14). The negative exponential function observed among freshwater fishes was influenced by the inclusion of desert (★; n=2) and non-teleost fishes (▲; n=3). When these species were removed from analyses, the decline of variation among freshwater Nearctic fishes with latitude was linear (r^2=0.61; $p<0.001$). Anadromous fishes exhibited a significant (r^2=0.35; $p<0.05$) decline of nucleotide diversity with latitude when non-teleosts (▲; n=2) were removed from analysis.


All freshwater fishes
Desert fishes removed
Desert fishes+non-teleosts removed
a) Freshwater fishes
Relative index of nucleotide diversity
0.08
0.06
0.04
0.02
0.00
30
35
40
45
50
55
Mean latitude of populations studied
Non-teleosts removed
b) Anadromous fishes
Relative index of nucleotide diversity
0.020
0.015
0.010
0.005
0.000
30
40
50
60
70
Mean latitude of populations studied

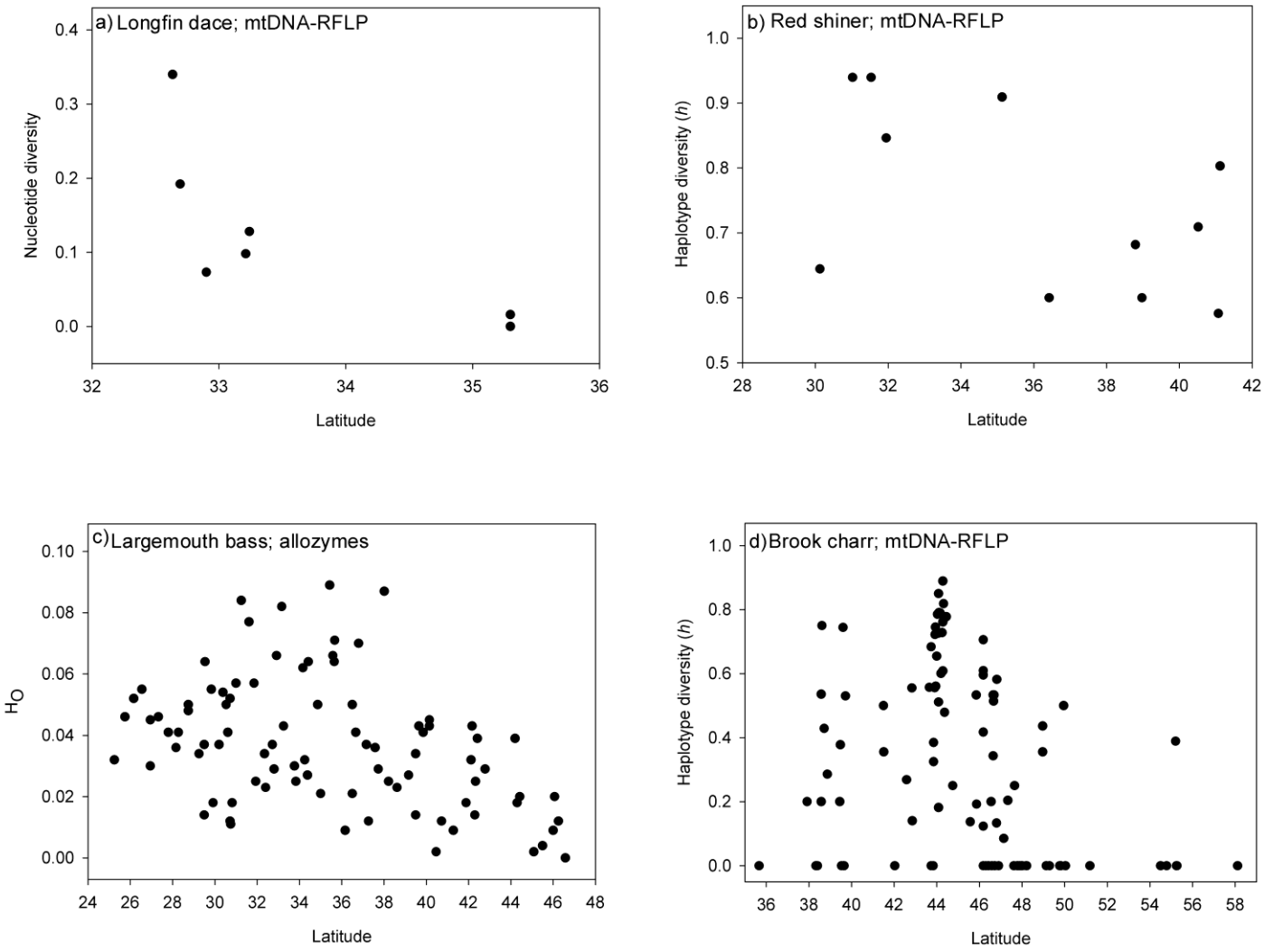
three populations, all of which were sampled from the southern portion of the species range (see Peterson et al. 2008), and is not likely to be as representative of the spatial distribution of genetic variation as the other measures which were estimated from across a more broad geographic region.

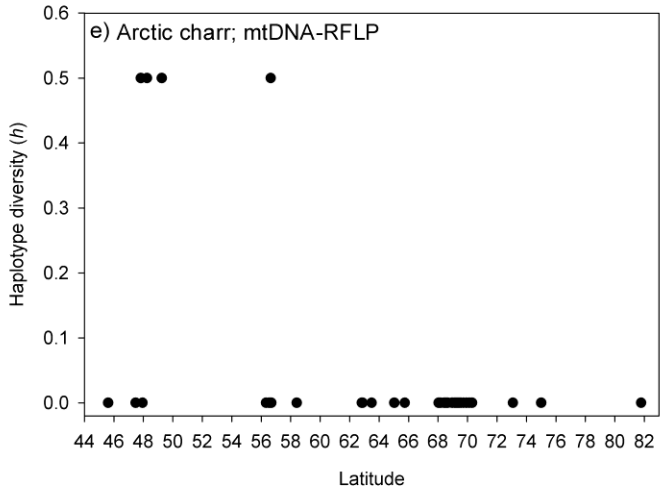
Overall, consistent evidence across marker types and analytical methods for negative correlations of intraspecific genetic diversity with latitude was observed for 24 Nearctic species. Of these species, linear regression analysis revealed significant ($p<0.05$) negative relationships between genetic diversity and latitude for 11 species (four freshwater, seven anadromous) (Table 2.2). However, a significant positive relationship was observed between intraspecific genetic diversity and latitude for lake whitefish, chinook salmon, bull trout, and bloater (Table 2.2). Linear regressions explained a greater proportion of variation in the decline of genetic diversity with latitude in anadromous fishes (mean $r^2=0.55\pm0.15$) (Figure 2.4e-o) than in freshwater species (mean $r^2=0.28\pm0.22$) (Figure 2.4a-d).

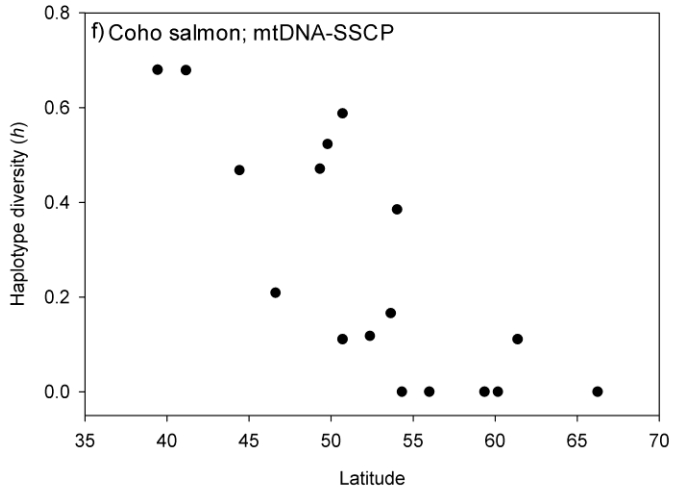
Non-linear declines of intraspecific genetic variation

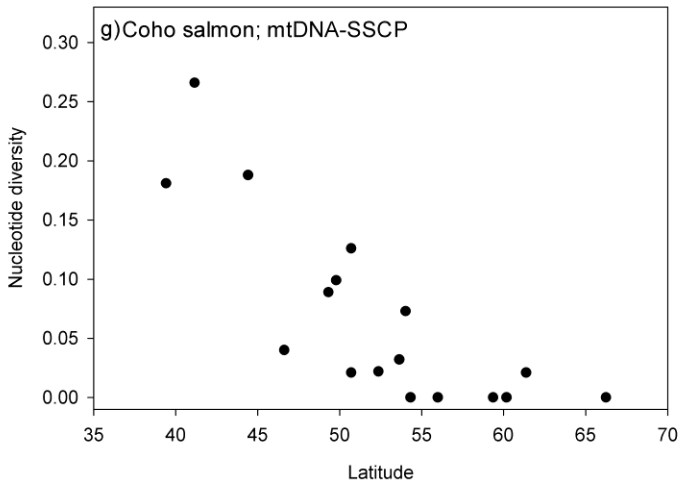
Five of the seven anadromous fishes distributed along the Atlantic coast considered in this study were among those species that exhibited significant ($p<0.05$) negative relationships of genetic diversity with latitude (i.e. Atlantic sturgeon, American shad, mummichog, striped bass, and rainbow smelt (glacial lineage 'A')) (Table 2.2). Negative correlations of intraspecific genetic diversity with latitude were not observed for Atlantic salmon or shortnose sturgeon. Visual inspection of the spatial distribution of genetic diversity for these five species (Figure 2.4h-o) suggested that linear regressions

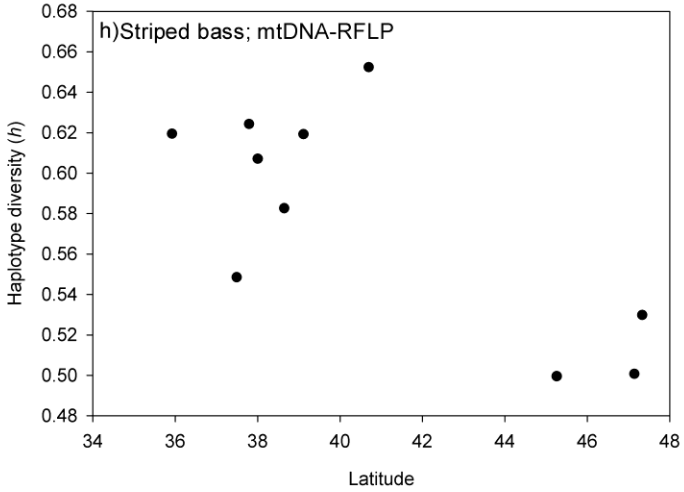
Figure 2.4 Spatial distributions of genetic diversity indices for freshwater and anadromous Nearctic fishes exhibiting significant ($p<0.05$) negative correlations of genetic variation with latitude: a) longfin dace (mtDNA-RFLP; π), b) red shiner (mtDNA-RFLP; *h*), c) largemouth bass (allozymes; H_O), d) brook charr (mtDNA-RFLP; *h*), e) Arctic charr (mtDNA-RFLP; *h*), f) coho salmon (mtDNA-SSCP; *h*), g) coho salmon (mtDNA-RFLP; π), h) striped bass (mtDNA-RFLP; *h*), i) Atlantic sturgeon (mtDNA-SSCP; *h*), j) rainbow smelt-glacial lineage 'A' (microsatellites; H_O), k) American shad (microsatellites; H_O), l) American shad (microsatellites; R), m) mummichog (allozymes; H_O), n) mummichog (microsatellites; H_O), o) mummichog (microsatellites; R).

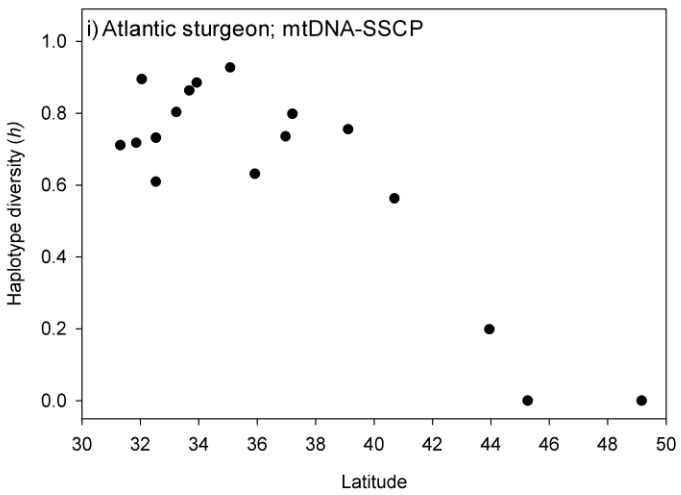

a) Longfin dace; mtDNA-RFLP
Nucleotide diversity
Latitude
b) Red shiner; mtDNA-RFLP
Haplotype diversity (h)
Latitude
c) Largemouth bass; allozymes
H_O
Latitude
d) Brook charr; mtDNA-RFLP
Haplotype diversity (h)
Latitude

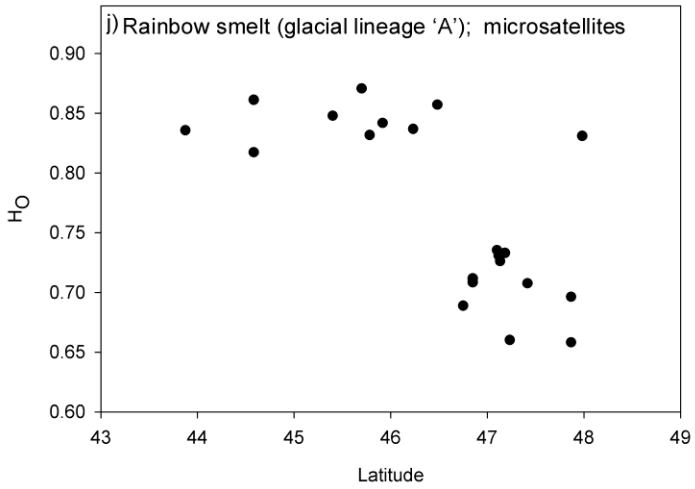

e) Arctic charr; mtDNA-RFLP
Haplotype diversity (h)
0.6
0.5
0.4
0.3
0.2
0.1
0.0
44 46 48 50 52 54 56 58 60 62 64 66 68 70 72 74 76 78 80 82
Latitude

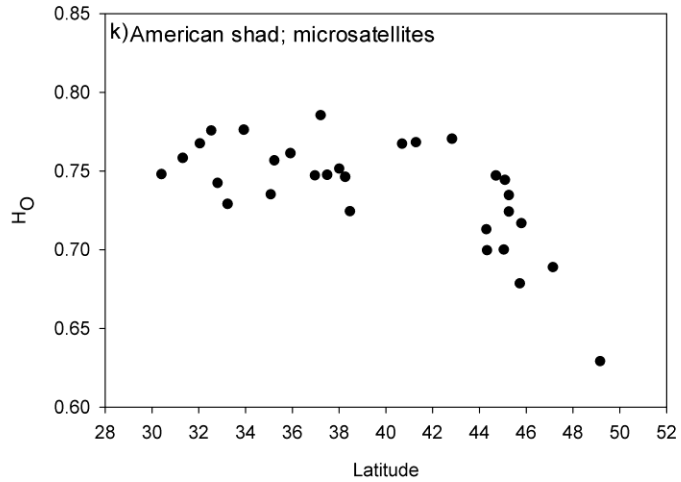

f) Coho salmon; mtDNA-SSCP
Haplotype diversity (h)
0.8
0.6
0.4
0.2
0.0
35 40 45 50 55 60 65 70
Latitude

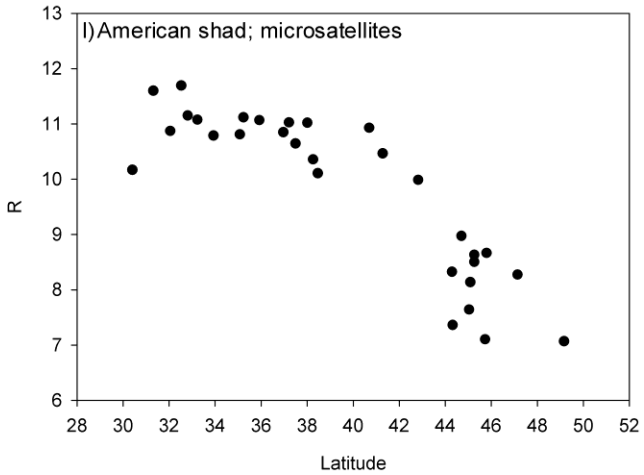

g)Coho salmon; mtDNA-SSCP
Nucleotide diversity
0.30
0.25
0.20
0.15
0.10
0.05
0.00
35 40 45 50 55 60 65 70
Latitude

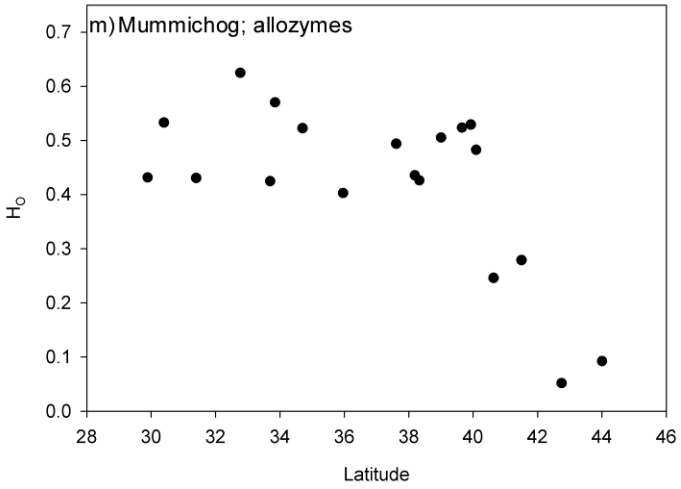

h)Striped bass; mtDNA-RFLP
Haplotype diversity (h)
0.68
0.66
0.64
0.62
0.60
0.58
0.56
0.54
0.52
0.50
0.48
34 36 38 40 42 44 46 48
Latitude

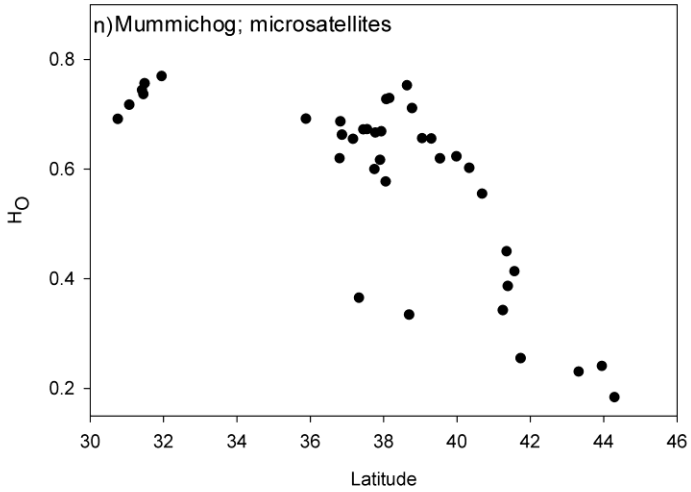

i) Atlantic sturgeon; mtDNA-SSCP
Haplotype diversity (h)
1.0
0.8
0.6
0.4
0.2
0.0
30
32
34
36
38
40
42
44
46
48
50
Latitude

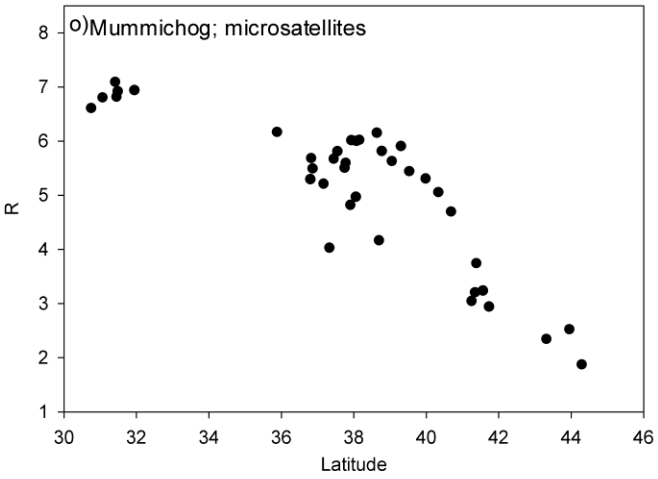

j) Rainbow smelt (glacial lineage 'A'); microsatellites
H_O
0.90
0.85
0.80
0.75
0.70
0.65
0.60
43
44
45
46
47
48
49
Latitude


k) American shad; microsatellites
H_O
0.85
0.80
0.75
0.70
0.65
0.60
28
30
32
34
36
38
40
42
44
46
48
50
52
Latitude


l) American shad; microsatellites
R
13
12
11
10
9
8
7
6
28
30
32
34
36
38
40
42
44
46
48
50
52
Latitude


m)Mummichog; allozymes
H_O
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
28
30
32
34
36
38
40
42
44
46
Latitude


n)Mummichog; microsatellites
H_O
0.2
0.4
0.6
0.8
30
32
34
36
38
40
42
44
46
Latitude


o)Mummichog; microsatellites
R
1
2
3
4
5
6
7
8
30
32
34
36
38
40
42
44
46
Latitude

may not accurately characterize latitudinal declines of intraspecific genetic diversity. Piecewise linear regressions conducted iteratively (1°N increments) across the range of latitudes surveyed for each species explained a greater proportion of variation (r^2=0.55-0.88) than linear regressions (r^2=0.38-0.76), and implied that latitudinal declines in genetic diversity were non-linear across broad spatial scales. The latitudinal breakpoint identified through piecewise linear regression varied slightly among species, and with the exception of rainbow smelt (glacial lineage 'A'), generally approximated the LGM (39.8°N±2.31) (Figure 2.5). The clusters of rainbow smelt populations represented in Figure 2.4j are not continuously distributed populations, but correspond to populations from Newfoundland (H_O<0.75) and the Canadian Maritime Provinces (H_O>0.80), respectively. Although the greatest proportion of variation in the decline of Atlantic sturgeon genetic diversity was observed at 35°N (r^2=0.88), this value was not appreciably reduced for alternative latitudinal breakpoints from 36°N-40°N (r^2=0.86-0.87) (Figure 2.5a). While non-significant spatial trends in genetic diversity were generally observed south of the latitudinal breakpoint (except mummichog allelic richness), significant ($p<0.05$) declines in variation were generally observed to the north of the latitudinal breakpoint (except striped bass haplotype diversity) (Figure 2.6). The decline in mummichog H_O (allozymes) above 40°N was marginally significant ($p=0.061$; Figure 2.6e), and supported by a strong negative correlation with latitude (-0.86).

Table 2.2 Regression statistics for significant ($p<0.05$) latitudinal clines of intraspecific genetic diversity for 15 Nearctic fishes.

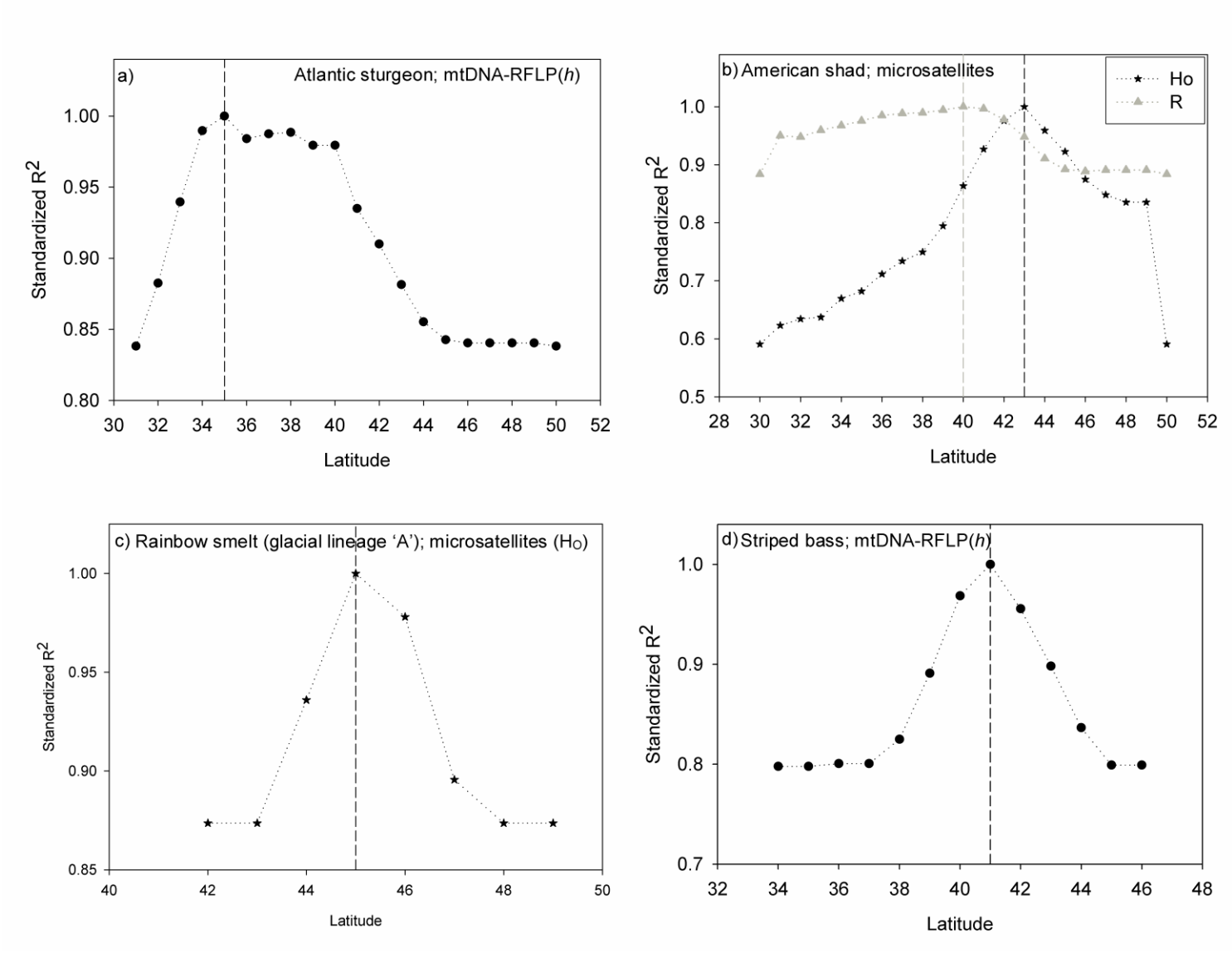
Species	Diversity measure	Marker	Analytical method	Pearson Correlation	r^2	P-value
Atlantic sturgeon	*h*	mtDNA	SSCP	-0.857[a]	0.735	<0.001
Longfin dace	π	mtDNA	RFLP	-0.754	0.568	<0.05
	H_O	Microsatellites		-0.648	0.420	<0.001
American shad	R	Microsatellites		-0.873	0.761	<0.001
Lake whitefish-normal (Mississippian refugium)	π	mtDNA	RFLP	0.496	0.246	<0.05
Lake whitefish-normal (refugia pooled)	π	mtDNA	RFLP	0.536	0.287	<0.001
Red shiner	*h*	mtDNA	RFLP	-0.572	0.327	<0.05
	h	Allozymes		-0.614	0.377	<0.01
	H_O	Microsatellites		-0.692	0.479	<0.001
Mummichog	R	Microsatellites		-0.828	0.685	<0.001

Table 2.2 *continued*

Species	Diversity measure	Marker	Analytical method	Pearson Correlation	r^2	P-value
Largemouth bass	H_O	Allozymes		-0.404	0.163	<0.001
Striped bass	*h*	mtDNA	RFLP	-0.746	0.556	<0.05
Coho salmon	*h*	mtDNA	SSCP	-0.797	0.635	<0.001
	π	mtDNA	SSCP	-0.811	0.657	<0.001
Chinook salmon	H_O	Microsatellites		0.898	0.807	<0.005
Rainbow smelt- glacial lineage 'A'	H_O	Microsatellites		-0.691	0.477	<0.001
Arctic charr	*h*	mtDNA	RFLP	-0.562	0.316	<0.001
Brook charr	*h*	mtDNA	RFLP	-0.277	0.077	<0.05
Bull trout	*h*	mtDNA	RFLP	0.737	0.543	<0.01
Bloater	π	mtDNA	RFLP	0.655	0.430	<0.01

a - D-loop

Figure 2.5 Piecewise linear regressions conducted iteratively for Atlantic coast anadromous Nearctic fishes to identify the latitude which explained the greatest proportion of variation in the spatial distribution of intraspecific genetic diversity. The short dashed line indicates the latitudinal breakpoint for the measure(s) of genetic diversity assessed among populations of a) Atlantic sturgeon, b) American shad, c) rainbow smelt (glacial lineage 'A'), d) striped bass, e) mummichog.


a)
Atlantic sturgeon; mtDNA-RFLP(h)
Standardized R2
Latitude
b) American shad; microsatellites
Ho
R
Standardized R2
Latitude
c) Rainbow smelt (glacial lineage 'A'); microsatellites (HO)
Standardized R2
Latitude
d) Striped bass; mtDNA-RFLP(h)
Standardized R2
Latitude

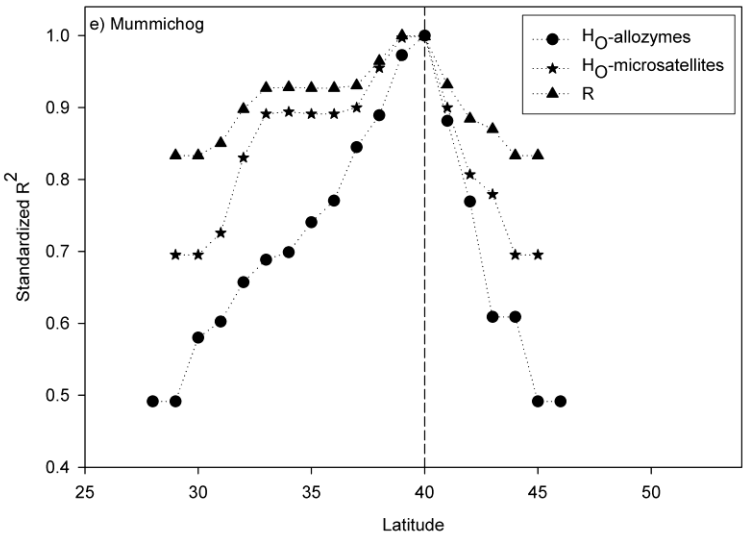
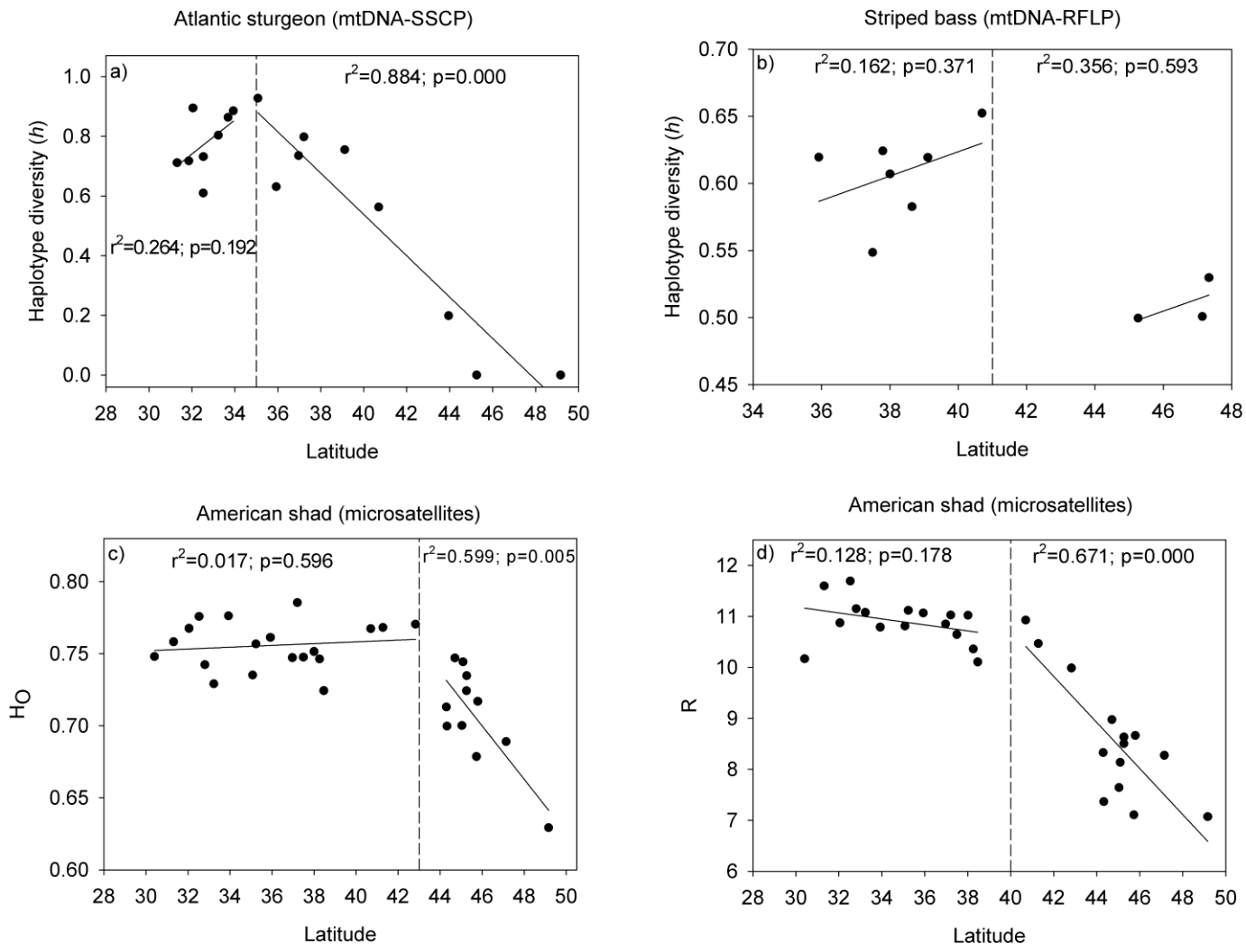
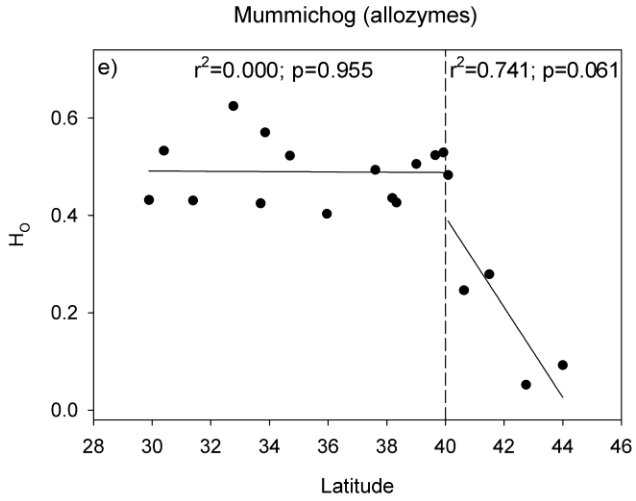
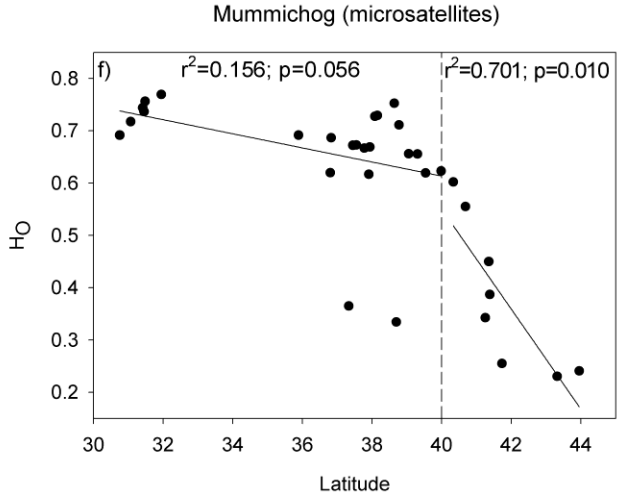
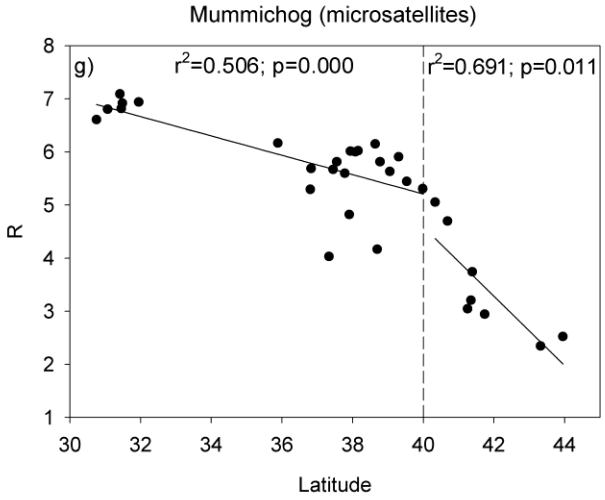

e) Mummichog
H_O-allozymes
H_O-microsatellites
R
Standardized R^2
Latitude
1.0
0.9
0.8
0.7
0.6
0.5
0.4
25
30
35
40
45
50

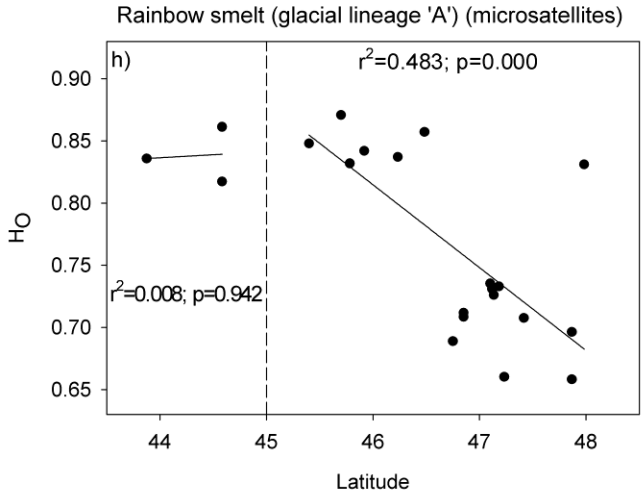
Figure 2.6 Piecewise linear regressions for five Atlantic coast Nearctic anadromous fishes demonstrated non-significant trends of intraspecific genetic diversity with latitude to the south of the latitudinal breakpoint, and significant ($p<0.05$) declines of genetic diversity with latitude to the north of the latitudinal breakpoint identified for each species. Panels correspond to the following species: a) Atlantic sturgeon (mtDNA-SSCP; *h*), b) striped bass (mtDNA-RFLP; *h*), c) American shad (microsatellites; H_O), d) American shad (microsatellites; R), e) mummichog (allozymes; H_O, f) mummichog (microsatellites; H_O), g) mummichog (microsatellites; R), h) rainbow smelt (glacial lineage 'A') (microsatellites; H_O).


Atlantic sturgeon (mtDNA-SSCP)
a)
r^2=0.884; p=0.000
r^2=0.264; p=0.192
Haplotype diversity (h)
Latitude
Striped bass (mtDNA-RFLP)
b)
r^2=0.162; p=0.371
r^2=0.356; p=0.593
Haplotype diversity (h)
Latitude
American shad (microsatellites)
c)
r^2=0.017; p=0.596
r^2=0.599; p=0.005
H_O
Latitude
American shad (microsatellites)
d)
r^2=0.128; p=0.178
r^2=0.671; p=0.000
R
Latitude


Mummichog (allozymes)
e)
r^2=0.000; p=0.955
r^2=0.741; p=0.061
H_O
Latitude


Mummichog (microsatellites)
f)
r^2=0.156; p=0.056
r^2=0.701; p=0.010
H_O
Latitude


Mummichog (microsatellites)
g)
r^2=0.506; p=0.000
r^2=0.691; p=0.011
R
Latitude


Rainbow smelt (glacial lineage 'A') (microsatellites)
h)
r^2=0.483; p=0.000
r^2=0.008; p=0.942
H_O
Latitude

Spatial distribution of genetic variation among species

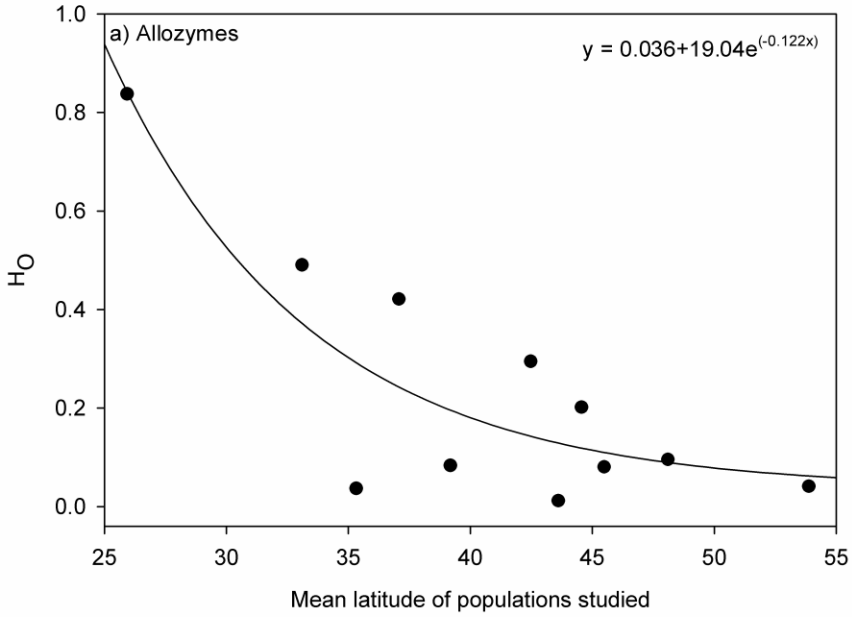
A strong negative correlation (r= -0.77) was observed among species between global estimates of intraspecific H_O (allozymes) and the mean latitude of the populations from which these estimates were obtained (Figure 2.7a). Negative exponential regression explained a substantially greater proportion of variation (r^2=0.75) in the decline of genetic variation among species than linear regression (r^2=0.59), and was significant ($p<0.01$). A negative correlation (r=-0.69) was also observed among species for mtDNA nucleotide diversity with mean study latitude. Linear regression revealed a significant ($p<0.05$) decline of nucleotide diversity with increasing mean latitude and explained a substantial proportion of variation (r^2=0.47) (Figure 2.7b).

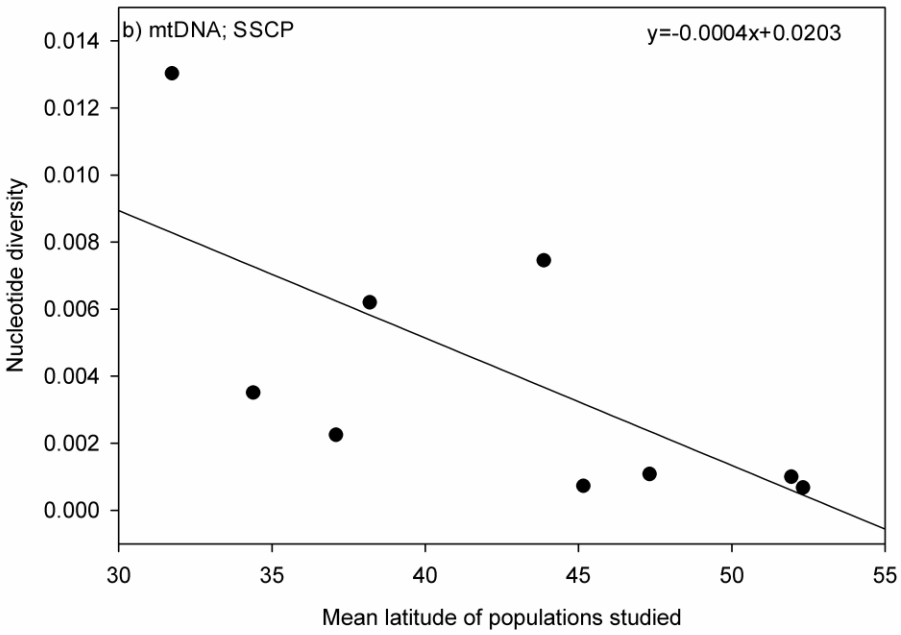
2.4 DISCUSSION

Spatial genetic variation within Nearctic fishes

This study confirms that latitude is an important factor in the spatial distribution of genetic variation within and among freshwater and anadromous Nearctic fishes. Negative correlations of intraspecific genetic diversity with latitude were observed for 82% of the species examined by Bernatchez and Wilson (1998) for which population level estimates of haplotype diversity could be calculated. The few instances of positive correlation observed (Table 2.1) may have been due to limited geographic scope of the source study, possible introgression among lineages derived from separate glacial refugia, or combinations thereof. The expanded data set generated from the literature survey revealed patterns broadly consistent with those found for the original source studies assessed by Bernatchez and Wilson (1998). Latitudinal declines of intraspecific variation

Figure 2.7 Latitudinal declines of genetic diversity among species from the literature survey. a) Negative exponential function best described the decline in H_O (allozymes). b) The decline in nucleotide diversity (SSCP) showed a similar trend.


a) Allozymes
y = 0.036+19.04e(-0.122x)
1.0
0.8
0.6
0.4
0.2
0.0
HO
25
30
35
40
45
50
55
Mean latitude of populations studied


b) mtDNA; SSCP
y=-0.0004x+0.0203
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
Nucleotide diversity
30
35
40
45
50
55
Mean latitude of populations studied

were observed across multiple marker types and analytical approaches for the majority of the species examined. However, positive correlations were observed for lake whitefish, chinook salmon, bull trout, and bloater, and may have been due to introgression among distinct glacial lineages, among other factors. Because this study classified populations to the glacial mtDNA lineage that comprised the greatest proportion of their genomes, introgression among lake whitefish lineages (e.g. between Mississippian and Beringian refugial lines) may have lead to elevated levels of diversity in northern regions (i.e. Arctic Red River, NWT; 67°N). The Beringian lineage has retained a high proportion of its diversity relative to other lineages (Bernatchez and Dodson 1991), and may explain the positive correlation observed between diversity and latitude when lineages were pooled. A similar explanation may apply for chinook salmon and bull trout. Populations of chinook salmon from British Columbia and Alaska are believed to have been derived from the Beringian refuge, while populations from California are thought to have been derived from a 'Pacific refuge' (McPhail and Lindsey 1970). Similarly, bull trout are believed to have been isolated in two separate refugia south of the Cordilleran ice sheet (i.e. the Chehalis and Columbia; Taylor et al. 1999). The positive correlation of intraspecific genetic diversity with latitude detected among bloater populations may be related to demographic fluctuations in population size experienced in the Great Lakes, and associated losses of genetic variation (Fave and Turgeon 2008).

Spatial genetic variation among Nearctic fishes

The prevailing influence of latitude on the distribution of intraspecific genetic variation argues for caution in comparisons among species, given that data within species

are rarely fully representative of species latitudinal ranges. When data are derived from only a portion of a species range, mean (or median) latitudinal values should correspond to the data sources, and not the species entire distribution. Species median latitudes reported by Bernatchez and Wilson (1998) fell outside of the mean (±95%) range of the populations assessed in the original source studies (from which estimates of global nucleotide diversity were obtained) for half of the species they considered, and hence did not provide an accurate geographic representation of estimates of intraspecific genetic variation. Although the difference between species median latitude and study mean latitude for some fishes was only slight (e.g. <1°, redear sunfish), it was considerably greater for other species (e.g. ~9°, lake trout). Although these latitudinal differences may have a small effect on the spatial characterization of genetic diversity estimates for any given species, the effects might amplify in comparisons across multiple species, causing some distortion of broad spatial scale patterns of genetic variation.

Bernatchez and Wilson (1998) observed a decline of genetic variation with latitude among species, and detected a latitudinal break at 46°N using median species latitudes with piecewise linear regression. However, I could not replicate those results with the data reported in Appendix I of their study and using piecewise linear regression for either species median latitude or study mean latitude. The source of this discrepancy remains uncertain. The spatial distribution of mtDNA genetic variation among the 36 Nearctic freshwater and anadromous fishes examined herein was characterized by a negative exponential function of the study mean latitude, and revealed greater variation among southern species than more northerly distributed species. A negative exponential decline of genetic variation was also observed among species for nuclear genetic markers

(H_O; allozymes), and implies that the spatial distribution of genetic variation among Nearctic fishes is best characterized as a negative exponential function of latitude.

Freshwater vs. anadromous species

Freshwater fishes were largely responsible for the exponential decline of nucleotide diversity observed among the 36 Nearctic fishes examined in this study. However, the negative exponential latitudinal decline observed among freshwater fishes was sensitive to the inclusion of desert species and non-teleost fishes. A variety of human induced changes to freshwater habitats in the southwestern U.S. (e.g. modification of flow regimes, desiccation of habitat, stream canalization/impoundment, introduction of exotic species) have combined to extirpate some desert fishes from much of their former range, and have substantially reduced population sizes and intraspecific levels of genetic variation (Dowling et al. 1996, Quattro et al. 1996, Tibbets and Dowling 1996). When two endangered desert fishes (i.e. Sonoran topminnow, razorback sucker) with low levels of nucleotide diversity were excluded from analysis, a highly significant linear decline of genetic variation with latitude was observed among remaining freshwater fishes. The low levels of nucleotide diversity exhibited by chondrostean and holostean fishes as observed by Bernatchez and Wilson (1998) may be due to a slower rate of mtDNA substitution associated with longer generation times and lower metabolic rates relative to teleosts (Krieger and Fuerst 2002, Avise et al. 1992). When lake sturgeon, bowfin, and Mississippi paddlefish were removed from analysis, linear regression was highly significant, and explained a greater proportion of variation (r^2=0.61) than either the negative exponential function among all freshwater fishes (r^2=0.30), or the linear decline identified through the exclusion of desert fishes alone (r^2=0.40) (Figure 2.3a).

Relative to freshwater species, anadromous fishes exhibited low levels of nucleotide diversity and comparatively little variation in the distribution of genetic variation across the same latitudinal range. Although a non-significant latitudinal effect was observed among all anadromous fishes (consistent with McDowall (1999)), the exclusion of non-teleosts revealed a significant linear latitudinal decline of genetic variation among remaining anadromous species; reflecting the pattern observed among freshwater teleosts, and consistent with the observations of Bernatchez and Wilson (1999).

As noted by McDowall (1999), comparatively low levels of genetic variation among anadromous fishes may be attributable to differences in patterns of post-glacial colonization as well as historical and contemporary connectivity among freshwater and anadromous species. Nucleotide diversity reflects both the frequency of various haplotypes as well as the evolutionary divergence among them. Successive population founder events during post-glacial colonization lead to sequential reductions of intraspecific genetic diversity with increasing latitude for both freshwater and anadromous Nearctic fishes. However, isostatic rebound following post-glacial colonization likely had a greater impact on levels of connectivity among populations of freshwater fishes than anadromous species. Isostatic rebound altered drainage patterns among freshwater systems, and eliminated connectivity among many populations, leading to greater sequence divergence among populations (Ward et al. 1994, Gyllensten 1985), and elevated levels of global nucleotide diversity within species. Connectivity among populations of anadromous fishes was probably not affected by isostatic rebound to the same extent as freshwater species because of their capacity for extensive marine

dispersal between reproductive events (but see Bernatchez and Dodson 1990, Bernatchez 1997). The comparatively low level of nucleotide diversity observed among anadromous fishes partially reflects the greater connectivity and reduced divergence observed among populations (Ward et al. 1994, Gyllensten 1985), and the effects of population founder events during post-glacial colonization.

The literature survey conducted herein revealed that freshwater fishes exhibited greater variation in genetic diversity estimates among populations than that observed among populations of anadromous fishes. This pattern may be related to differences in post-glacial colonization processes among inland vs. coastal regions, and contemporary levels of connectivity. Following glacial recession, anadromous fishes may have gained access to new habitats as they became sequentially suitable for colonization (i.e. a stepping stone process; Ibrahim et al. 1996), leading to successive reductions in intraspecific genetic variation with latitude. Because of their capacity for marine dispersal, contemporary connectivity among populations of anadromous fishes is probably maintained to a greater extent than freshwater fishes. The process of post-glacial colonization of freshwater systems among inland regions may have been more complex, with initial access to habitats dependent on the removal of ice dams and glacial tills (Curry 2007, Pielou 1991), and connectivity among populations dependent on the establishment, alteration, and elimination of post-glacial colonization routes by isostatic rebound. Stream capture among drainages may have brought into secondary contact populations which experienced different demographic histories in isolated glacial refugia (e.g. lake whitefish, lake trout), confounding latitudinal patterns, and increasing the amount of variation in the spatial distribution of intraspecific genetic diversity. A greater

understanding of the spatial distribution of intraspecific genetic diversity among freshwater fishes may be gained through examination of the 'colonization distance' of populations from their glacial refugia (or the glacial margin) during the LGM. However, such distances may be difficult to assess because of the effects of isostatic rebound and the absence of a comprehensive chronology of major drainage connections during post-glacial colonization.

While post-glacial colonization may have played a dominant role in the greater variation of intraspecific genetic diversity observed among freshwater fishes, anthropogenic factors may also have influenced contemporary patterns. Several of the freshwater species assessed support valuable recreational fisheries (e.g. largemouth bass, smallmouth bass, brook charr, lake trout, walleye etc.), some of which may have been subjected to artificial enhancement through stocking practices, including stock transfers among genetically distinct populations.

Non-linear declines of intraspecific genetic variation

The literature survey conducted herein revealed consistent evidence for non-linear declines of intraspecific genetic diversity with latitude across broad spatial scales for five of seven Atlantic anadromous fishes. Among these species, the greatest proportion of variation in piecewise regression was consistently explained by a latitudinal breakpoint which approximated the LGM (40°N), with the exception of rainbow smelt. Evidence for significant declines of intraspecific genetic diversity indices among populations north of this latitude was observed for six of eight diversity index-species comparisons. Neither shortnose sturgeon nor Atlantic salmon exhibited negative correlations of intraspecific

genetic diversity with latitude. However, northern populations of shortnose sturgeon are believed to have persisted in more than one glacial refugium during the LGM (Grunwald et al. 2002) and, similar to chinook salmon and bull trout, may have experienced introgression between distinct refugial lineages, obscuring spatial patterns of genetic variation. Contemporary populations of Atlantic salmon do not span the LGM. Given the relevance of the LGM to the spatial distribution of genetic variation for other Atlantic anadromous fishes, negative correlations of genetic diversity with latitude for Atlantic salmon might not be expected to be as strong or consistent across marker types. Alternatively, anthropogenic factors may have influenced the contemporary distribution of intraspecific variation for Atlantic salmon (e.g. substantial declines in abundance inversely correlated with latitude, or supportive breeding including stock transfer among genetically distinct populations). Overall, these results are consistent with the successive loss of genetic variation following sequential post-glacial colonization of Atlantic coastal rivers by anadromous fishes, and suggest that the portion of the Atlantic coast north of the LGM may constitute an empirical example of a one-dimensional stepping stone model of post-glacial colonization.

2.5 SUMMARY

This study confirms that latitude is an important determinant of the spatial distribution of genetic variation for many Nearctic fishes; latitudinal declines in genetic diversity among species are mirrored by spatial declines within species. Thus, meaningful interpretation of broad spatial scale inter-specific patterns of genetic variation requires geographically appropriate representation of intraspecific genetic diversity. Many of the original source studies considered by Bernatchez and Wilson (1998)

examined only a portion of a species range (e.g. Atlantic sturgeon). Results from this study suggest that the prior use of species median latitudes did not accurately represent the spatial distribution of the populations from which estimates of genetic variation were obtained, and may have introduced a bias that obscured broad spatial scale patterns (i.e. latitudinal declines) of genetic variation. This study demonstrates that latitudinal declines in intraspecific genetic diversity for species with broad geographic ranges (e.g. Atlantic anadromous fishes) are non-linear, and that a thorough understanding of the spatial distribution of intraspecific genetic variation requires a range-wide approach.

Latitudinal declines of genetic diversity among Nearctic fishes are best characterized by a negative exponential function. However, this pattern was largely driven by the exponential decline of variation observed among freshwater fishes, which in turn was influenced by the inclusion of desert species and non-teleost fishes. Exclusion of these species from analysis revealed a linear latitudinal decline of genetic variation among remaining freshwater teleosts. Although a non-significant latitudinal effect was observed among all anadromous fishes, exclusion of non-teleosts from analysis revealed a significant linear decline of genetic variation with latitude among remaining anadromous fishes, consistent with the pattern observed among freshwater fishes. Dissimilar amounts of genetic variation among freshwater and anadromous fishes likely reflect differential patterns of post-glacial colonization, as well as historical and contemporary connectivity among populations.

Comparisons among Atlantic anadromous fishes revealed consistent evidence for the influence of the LGM on contemporary spatial distributions of genetic variation, and suggests that the portion of the Atlantic coast north of the LGM may provide an empirical

example of a one-dimensional stepping stone model of post-glacial colonization. While historical demographic factors associated with climatic change have probably played a dominant role in the distribution of genetic diversity for Nearctic fishes, interpretations of spatial genetic variation also require consideration of the relative influences of contemporary microevolutionary processes and anthropogenic factors.

CHAPTER 3

TAKING STOCK: DEFINING POPULATIONS OF AMERICAN SHAD IN CANADA USING NEUTRAL GENETIC MARKERS

3.1 INTRODUCTION

Knowledge of the spatial scale of population structure is important to effective resource management. The identification of population boundaries is a fundamental prerequisite to the designation and prioritization of conservation units, for the protection of adaptive genetic diversity, and the preservation of species evolutionary potential (Allendorf 1986). Population designations which emphasize reproductive cohesion (i.e. evolutionary paradigm; Waples and Gaggiotti 2006) factor prominently in defining population structure, as highly polymorphic molecular markers and advanced statistical approaches are increasingly sought to inform conservation and management decisions (Crandall et al. 2000; Allendorf et al. 2004). For species where segregation into demes is certain only during reproduction, the essence of the evolutionary paradigm may manifest during a comparatively narrow phase of their life history. Prior to first reproduction and between reproductive events, individuals may occur in mixed assemblages, perhaps during extensive annual migrations. In such scenarios, population delineation based on demographic interactions (i.e. ecological paradigm; Waples and Gaggiotti 2006) may incorrectly suggest the existence of a single population, whereas molecular signatures inherent to reproductive connectivity may provide more accurate resolution. These considerations apply to many species, but are particularly relevant to the management of anadromous fishes.

American shad (*Alosa sapidissima* Wilson, 1811) is an anadromous clupeid indigenous to the Atlantic coast of North America, and spawns in rivers from Florida to Quebec (Walburg and Nichols 1967). South of Cape Hatteras, NC, American shad are semelparous, while to the north iteroparity increases with latitude (Leggett and Carscadden 1978). Juveniles emigrate from natal rivers and shoal as mixed assemblages along the coast until maturity (Talbot and Sykes 1958; Dadswell et al. 1987). Reproductive isolation among spawning runs is maintained by a combination of philopatry (Hendricks et al. 2002) and spawning site fidelity (Melvin et al. 1986). Molecular studies (e.g. Bentzen et al. 1989; Waters et al. 2000) have revealed genetic differences among shad spawning runs, and provide powerful management tools for resolving population structure.

American shad are of increasing conservation concern as range-wide declines in abundance attributable to anthropogenic factors (Bilkovic et al. 2002a) have reduced the number of contemporary spawning runs to fewer than half of their historic level (Limburg et al. 2003), and have resulted in moratoria being imposed upon once profitable fisheries (ASMFC 1999 as in Bilkovic et al. 2002b). Previous molecular studies of shad have predominantly focused on spawning runs in the United States, and revealed that individual rivers support genetically distinguishable spawning runs (Bentzen et al. 1989). However, only five of the more than 20 suspected Canadian river populations (Chaput and Bradford 2003 and references therein) have ever been examined (i.e. St. Lawrence, Miramichi, Annapolis, Shubenacadie, Saint John) (e.g. Bentzen et al. 1989; Epifanio et al. 1995; Waldman et al. 1996), and previous molecular studies mainly employed mitochondrial DNA (mtDNA) to resolve population structure. While highly polymorphic

microsatellite markers potentially offer greater resolution of population structure, prior microsatellite based studies of shad were limited to few loci and few populations, and did not include Canadian spawning runs (Brown et al. 2000; Waters et al. 2000). Knowledge of the scale of population structure among Canadian spawning runs remains unknown, but must be understood for effective management and conservation of shad genetic diversity.

I use microsatellites to examine the spatial partitioning of American shad genetic variation among reproductive demes within the Canadian portion of the species range, and assess the temporal stability of population structure at multiple spatial scales. I identify the number of genetically distinguishable shad populations in Canada; information that is directly applicable for management and conservation.

3.2 Materials and Methods

3.2.1 Sample Collections

Fisheries data have revealed a high degree of uncertainty as to the spatial scale of American shad population structure in Canada (Chaput and Bradford 2003). To capture as much of the potential genetic variation as possible, I sampled river specific commercial fisheries with a multi-decadal scale of exploitation, and supplemented these with targeted search efforts for spawning runs not subject to commercial exploitation. Sampling from 2003-2006 employed a combination of trapnets, gillnets, seines, and angling, and provided 2,555 fin clips for 24 sites from 12 rivers (Table 3.1; Figure 3.1).

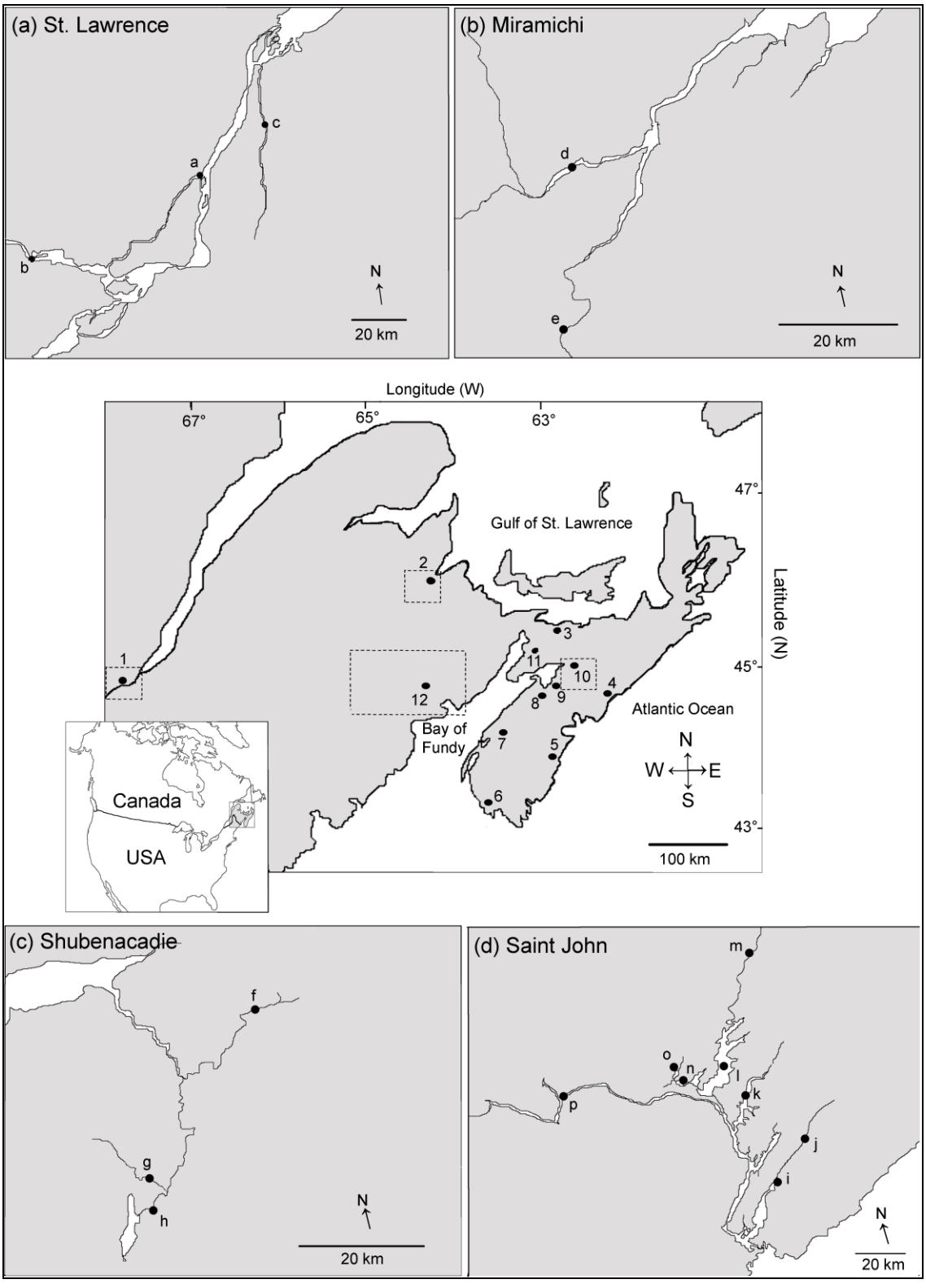
Table 3.1 Sampling locations and sample sizes (N) for American shad collected from 2003-2006 from Atlantic Canada.

Abbr.	River	Tributary	2003	2004	2005	2006	Total
STL	St. Lawrence, QC	Rivière des Prairies		106	102		208
		Ottawa River			99		99
		Richelieu River		13	22		35
MR	Miramichi River, NB	Miramichi River		152			152
		NW Miramichi River		15	99		114
		SW Miramichi River		35	97	33	165
SHU	Shubenacadie River, NS	Nine Mile River		101	103	130	334
		Shubenacadie River		84	100	100	284
		Stewiacke River		102	43		145
SJR	Saint John River, NB	French Lake				48	48
		Indian Lake				41	41
		French/Indian lakes combined	32				32
		Grand Lake		28	30	45	103

Table 3.1 *continued*

Abbr.	River	Tributary	2003	2004	2005	2006	Total
SJR	Saint John River, NB	Hammond River				68	68
		Kennebecasis River	30				30
		Mactaquac Dam			99		99
		Salmon River				20	20
ANN	Annapolis River, NS	Washademoak Lake	33				33
				111			111
AR	Annis River, NS			10		7	17
GR	Gaspereau River, NS			85		35	120
KR	Kennetcook River, NS				50		50
LH	LaHave River, NS			12	15	41	68
MUS	Musquodoboit River, NS			12	20	13	45
RH	River Hebert, NS			23	22		45
RP	River Phillip, NS			16	22	51	89

Figure 3.1 Map of Atlantic Canada (centre) displaying sampling locations for collections of American shad: 1-St.Lawrence River (a-Rivière des Prairies; b-Ottawa River; c-Richelieu River); 2-Miramichi River (d-Northwest Miramichi River; e-Southwest Miramichi River); 3-River Phillip; 4-Musquoboboit River; 5-LaHave River; 6-Annis River; 7-Annapolis River; 8-Gaspereau River; 9-Kennetcook River; 10-Shubenacadie River (f-Stewiacke River; g-Nine Mile River; h-Shubenacadie River); 11-River Hebert; 12-Saint John River (i-Hammond River; j-Kennebecasis River; k-Washademoak Lake; l-Grand Lake; m-Salmon River; n-French Lake; o-Indian Lake; p-Mactaquac Dam).


(a) St. Lawrence
a
b
c
N
20 km
(b) Miramichi
d
e
N
20 km
Longitude (W)
67°
65°
63°
47°
45°
43°
Latitude (N)
Gulf of St. Lawrence
1
2
3
4
5
6
7
8
9
10
11
12
Bay of Fundy
Atlantic Ocean
N
W
E
S
100 km
Canada
USA
(c) Shubenacadie
f
g
h
N
20 km
(d) Saint John
m
o
n
l
k
p
j
i
N
20 km

Collections (i.e. samples obtained from sites within years) were replicated for a subset of sites over successive years to quantify temporal genetic variation. Samples consisted of adults captured in rivers above the influx of saltwater, with the following exceptions: i) the Kennetcook River collection was composed primarily of juvenile specimens; ii) River Phillip and Miramichi River (2004-2005) collections consisted of adults obtained in tidal freshwater (although the 2006 Miramichi River collection was sampled from above the head of tide in a major tributary; the southwest Miramichi River). Fin clips were preserved in 95% ethanol until DNA extraction. Within year collections from River Phillip, River Hebert, Annis, LaHave, and Musquodoboit rivers failed to meet target sample sizes ($n \geq 50$); consequently, samples from these rivers were pooled across years following analysis of molecular variance (AMOVA) which revealed non-significant ($p > 0.05$) genetic variation among years within rivers.

3.2.2 LABORATORY PROTOCOLS

DNA isolation followed the method of Elphinstone et al. (2003), modified for automation in 96 well format using a Perkin Elmer MPII liquid handler. Specimens were genotyped across 14 polymorphic microsatellite loci developed for *A. sapidissima* ((*AsaB020, AsaD029, AsaD042, AsaD429*; Julian and Bartron 2007) (*Asa16*; Brown et al. 2000) (*Asa2, Asa4, Asa8*; Waters et al. 2000)), *A. fallax* (*Af6, Af13;* Faria et al. 2004), *A. alosa (Af20, Aa14, Aa16*; Faria et al. 2004), and *A. pseudoharengus* (*Aps2A;* Bentzen and Paterson 2005). Polymerase chain reaction (PCR) amplifications were conducted in 5μL volumes consisting of 4mM $MgS0_4$, 2X *TSG* DNA polymerase reaction buffer, 0.4mM dNTP mix, forward (fluorescently labeled) and reverse primers (1μM for IR700/IR800 labeled primers; 2μM for HEX/FAM labeled primers), 0.05U *TSG* DNA polymerase, and

10-50ng DNA. PCR was conducted using either MJ Research PTC-225 Thermal Cycler or Eppendorf mastercyclers. Amplification cycles for all loci consisted of an initial denature step at 95°C for 3 min, followed by 30 cycles of denature for 30 s at 95°C, annealing for 30 s at locus specific temperatures, extension for 30 s at 72°C, and a final extension at 72°C for 3 min. PCR products were visualized on 6% denaturing polyacrylamide gels on either an FMBIO fluorescent imaging system (Hitachi Software Engineering Co. Ltd.) or LiCor IR2 DNA analyzer (LiCor Inc.) depending on the fluorescent label of the forward primer. Alleles were scored visually by reference to a molecular weight size standard constructed from pUC18 derived PCR fragments. Gel images were scored with reference to positive controls, negative controls, and redundant samples to ensure consistency and reproducibility of genotypes across collections and loci.

3.2.3 DATA ANALYSIS

Data conformance to model assumptions

The presence of genotyping artefacts was evaluated using Microchecker v.2.2.3 (van Oosterhout et al. 2004). Tests for departures from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were performed with GENEPOP v. 4.0.6 (Rousset 2007) using default parameters for all tests. Sequential Bonferroni adjustments were used to judge significance levels for all simultaneous tests (Holm 1979; Rice 1989). Selective neutrality of the microsatellite markers used in this study was evaluated by assessing relative variance in repeat number (lnRV) and heterozygosity (lnRH) (Schlötterer 2002; Schlötterer and Dieringer 2005). Because inclusion of juveniles may inflate population differentiation estimates if they are the progeny of few parents, the

Kennetcook collection was tested for the presence of kin groups following the method of Smith et al. (2001) as implemented in PEDIGREE (http://herbinger.biology.dal.ca:5080/Pedigree/).

Genetic diversity

For each collection the number of alleles per locus (N_a), observed heterozygosity (H_O), an unbiased estimate of heterozygosity (H_E; Nei 1978), and F_{IS} (Weir and Cockerham 1984) were calculated using the program GENETIX v. 4.05 (Belkhir et al. 2004). Allelic richness (R) per locus and collection were estimated using Fstat 2.9.3.2 (Goudet 1995; 2001) standardized to a minimum sample size of 32 individuals (Leberg 2002). Miramichi River (2006) and Annis River collections were excluded from allelic richness estimation due to sample size limitations (Table 3.1).

Genetic differentiation

The statistical power and realized α-error for assessing the null hypothesis of genetic homogeneity within and among rivers was assessed using POWSIM (Ryman and Palm 2006). Allelic heterogeneity within and among rivers was assessed via genic tests conducted with GENEPOP v.4.0.6 (Rousset 2007), and with default parameters for all tests. Tests were combined across loci or collections using Fisher's method. Hierarchical AMOVA was conducted to partition components of genetic variation among collections, among years within collections, and among individuals within collections, using a permutation procedure (10,000 iterations) (Arlequin 3.1; Excoffier et al. 2005).

Genetic differentiation was estimated by computing overall and pairwise F_{ST} (θ; Weir and Cockerham 1984) using FSTAT 2.9.3.2(Goudet 1995; 2001). The effects of

variation in genetic diversity on genetic differentiation (Hedrick 2005) were controlled for by calculating standardized estimates of differentiation (F'_{ST}) using RECODEDATA v. 0.1 (Meirmans 2006), which recoded the data so that there were no shared alleles among rivers. Recoded data were then used to estimate $F_{ST(max)}$ for each pairwise comparison using FSTAT (Goudet 1995; 2001). Standardized pairwise estimates of differentiation were then calculated as $F'_{ST} = F_{ST} / F_{ST(max)}$ following Hedrick (2005).

Relationships among rivers

Genetic affinities among rivers were examined with unrooted neighbour-joining (NJ) trees based on chord distance (D_C; Cavalli-Sforza and Edwards 1967) and bootstrapped over loci (10,000 replicates) computed with POPULATIONS v. 1.2.30 (Langella 1999) and viewed with TREEVIEW (Page 1996).

Population structure

Two Bayesian model-based clustering methods, implemented in STRUCTURE v. 2.2 (Pritchard et al. 2000; Falush et al. 2003) and BAPS v. 5.1 (Corander et al. 2006), were used concurrently to infer the number of genetically homogenous clusters (populations), as advocated by Latch et al. (2006). For STRUCTURE analysis, a burn-in of 50,000 replicates was followed by 250,000 replicates of the Markov chain Monte Carlo (MCMC) simulation, employing the admixture model and correlated allele frequencies among populations. To assess convergence of the log probability of data Pr(X|K) among rivers, three iterations of this parameter set were performed for K (number of clusters) from 1-15, allowing an estimation of the most likely number of

clusters. Both the plateau of likelihood values (Pritchard et al. 2000) and ΔK (i.e. second order rate of change between successive K values; Evanno et al. 2005) were estimated.

For the BAPS analysis, the mixture model was first applied to cluster groups of individuals based on their multilocus genotypes. Three iterations of K (1-15) were conducted among rivers to determine the optimal number of genetically homogeneous groups. Admixture analysis (Corander and Marttinen 2006) was then applied to estimate individual admixture proportions with regards to the most likely number of K clusters identified (Corander and Marttinen 2006), and were visualized using DISTRUCT v. 1.1 (Rosenberg 2004).

BARRIER analysis

BARRIER v. 2.2 (Manni et al. 2004) was used to define geographic discontinuities in gene flow among rivers. Barriers were first computed using a multilocus F_{ST} matrix along with geographic coordinates for each population to represent genetic differentiation among populations. F_{ST} matrices for each of the 13 loci were then used separately to identify 'consensus barriers' (i.e. ≥ 7 loci) to assess the robustness of the barriers identified using multilocus F_{ST}. Connectivity among rivers tributary to the Bay of Fundy was measured in two ways: (i) constrained to reflect the counterclockwise migration route postulated by Dadswell et al. (1983; 1987), and (ii) full connectivity among rivers, following the most direct route and incorporating trans-Bay of Fundy measures. Although the counterclockwise migration route was observed specifically among non-spawning and post-spawning adults engaged in a feeding migration

(Dadswell et al. 1983; 1987; Melvin et al. 1986), I considered this pattern to be feasible among pre-spawning adults as well.

Isolation by distance

An analysis of isolation by distance (IBD) was conducted among rivers to test for correlations between geographic distance and genetic differentiation using 10,000 permutations of the Mantel test as implemented in the web-service program IBDWS v. 3.15 (http://www.ibdws.sdsu.edu) (Jensen et al. 2005). Pairwise F_{ST} values were linearized as ($F_{ST}/(1-F_{ST})$) following Rousset (1997). Geographic distance between rivers was measured as the most direct route following within 5 km of the coastline using Google Earth 2006, following Bradbury et al. (2006; 2008). This specific distance was chosen as pre-spawning adults are known to migrate within a few kilometers of the coastline during their spawning migration (Leggett and Whitney 1972; Neves and Depres, 1979). Distances between Bay of Fundy rivers were measured following the two migration schemes outlined for BARRIER analysis (see above).

3.3 RESULTS

3.3.1 DATA CONFORMANCE TO MODEL ASSUMPTIONS

Locus *Aa16* was removed from analysis because null alleles were detected for >50% of collections examined (Microchecker). Remaining loci were retained for analyses, as evidence for null alleles were sporadically distributed among loci and collections. Exact tests revealed that genotypic frequencies were largely in accordance with HWE ($p>0.05$; sequential Bonferroni correction for 21 comparisons). However, departures from HWE remained for 22 locus-collection comparisons after sequential

Bonferroni adjustment, 21 of which consisted of shad from multiple collection sites within rivers (i.e. St. Lawrence, Miramichi, Shubenacadie, Saint John River; Table 3.1). Global score (U) tests revealed significant ($p<0.05$) heterozygote deficiencies among collection sites within these drainages, and suggested the potential for population sub-structuring. Exact tests of LD revealed that loci were physically unlinked and statistically independent ($p>0.05$; sequential Bonferroni correction for 1638 comparisons). Relative variance in repeat number (lnRV) and heterozygosity (lnRH) failed to detect outlier loci (i.e. markers were normally distributed; data not shown), and provided no evidence of non-neutrality among the loci examined. The juvenile collection from the Kennetcook River (n=50; Table 3.1) was partitioned into 28 kin groups, and was not the product of a small number of reproducing adults. Therefore, the allele frequency distributions were considered representative of an adult spawning population, and this collection was retained in analyses.

Shad are known to venture into non-natal estuaries (Leggett 1977a) and to move among adjacent regions just prior to upstream spawning migration (Dodson and Leggett 1974). Although some collections were obtained from tidal freshwater, population admixture is not likely a confounding factor, as non-significant heterozygote deficiency ($p\geq0.05$; sequential Bonferroni correction for 273 comparisons) was detected for all collections (data not shown). Although this may constitute a relatively weak test (i.e. individuals from non-natal rivers would need to be both numerous in the sample and substantially genetically differentiated from the target population for population admixture to be detectable as a heterozygote deficiency), I also failed to detect departure from HWE for any locus of the River Phillip collection, which suggested that this

collection is representative of a single, randomly mating population. The HWE departure observed for the Miramichi collections was likely due to the significant genic differentiation observed between tributaries within this drainage (see below).

3.3.2 GENETIC DIVERSITY

The degree of microsatellite polymorphism among the 12 rivers varied greatly depending on the locus and collection considered (Appendix 3). The number of alleles per locus ranged from four (*Af20*) to 39 (*Aa14*), with 11 loci exhibiting ≥ 12 alleles. Unbiased heterozygosity varied between 0.636 (STL05) and 0.758 (MUS), and allelic richness ranged from 6.64 (STL05) to 8.90 (MUS) (Appendix 3).

3.3.3 GENETIC DIFFERENTIATION

An assessment of statistical power indicated that my data was sufficient to detect weak differentiation both among and within rivers, while maintaining the realized α-error near the intended level (0.05) for tests of genetic homogeneity. The probability of obtaining a significant ($p<0.05$) result in contingency tests among rivers with an F_{ST} of 0.001 was 0.998 (χ^2), while that within rivers with an F_{ST} of 0.0025 was ≥0.860 (χ^2).

Exact tests revealed evidence of genetic heterogeneity within drainages. Significant ($p<0.05$) genic differentiation was observed among tributaries within each of the St.Lawrence, Miramichi, Shubenacadie, and Saint John River systems. However, differentiation was not temporally stable, as highly significant ($p<0.001$) heterogeneity was detected across years within tributaries, and suggested temporal variation in allele frequencies among tributaries within drainages (Table 3.2). While this pattern is perhaps best demonstrated for the Shubenacadie drainage (i.e. adequate representation of all

Table 3.2 Analysis of molecular variance among tributaries within rivers, among rivers, and among clusters identified using Bayesian clustering analyses.

Rivers	Variance component	d.f.	Sum of squares	% total variance	F-statistic	p-value
St. Lawrence	Among tributaries	1	8.14	0.21	0.002	0.125
	Among years within tributaries	2	19.03	1.28	0.013	<0.001
	Among individuals within tributaries	382	1691.19	98.50	0.015	<0.001
Miramichi	Among tributaries	1	7.44	-0.66	-0.007	0.617
	Among years within tributaries	2	15.40	1.27	0.013	<0.001
	Among individuals within tributaries	424	1982.79	99.39	0.006	<0.001

Table 3.2 *continued*

Rivers	Variance component	d.f.	Sum of squares	% total variance	F-statistic	p-value
Shubenacadie	Among tributaries	2	9.70	-0.09	0.000	0.964
	Among years within tributaries	5	34.08	0.24	0.002	<0.001
	Among individuals within tributaries	1482	6973.07	99.86	0.001	<0.001
Saint John	Among tributaries	7	62.85	-0.53	-0.005	0.685
	Among years within tributaries	3	26.67	1.75	0.017	<0.001
	Among individuals within tributaries	857	3839.38	98.78	0.012	<0.001

Table 3.2 *continued*

Rivers	Variance component	d.f.	Sum of squares	% total variance	F-statistic	p-value
All	Among rivers	11	761.45	4.12	0.041	<0.001
	Among years within rivers	17	75.19	0.05	0.001	0.097
	Among individuals within rivers	4907	20049.00	95.83	0.042	<0.001
Clusters	Among clusters	6	703.62	3.94	0.039	<0.001
	Among rivers within clusters	22	133.02	0.32	0.003	<0.001
	Among individuals within rivers	4907	20049.00	95.74	0.043	<0.001

tributaries across all years; Table 3.1), highly significant ($p<0.001$) temporal heterogeneity was also observed for Rivière des Prairies (St. Lawrence River), Southwest Miramichi River (Miramichi River), and Grand Lake (Saint John River). Hierarchical AMOVA revealed a non-significant ($p>0.05$) component of variation partitioned among tributaries within each of the St. Lawrence, Miramichi, Shubenacadie, and Saint John River systems (Table 3.2), and STRUCTURE v.2.2 identified a single cluster as the most likely number of genetically homogenous groups within each of these drainages (data not shown). Re-analyses of these data incorporating information on sampling site location (as implemented in STRUCTURE v.2.3; Hubisz et al. 2009) did not alter these results.

Pairwise comparisons revealed significant genic differentiation between all rivers ($p<0.05$ for Musquodoboit vs. LaHave Rivers; $p<0.001$ for all other comparisons); a multilocus standardized estimate of global differentiation (F'_{ST}) was 0.16 (F_{ST}=0.043). Pairwise standardized estimates of differentiation among rivers ranged from 0.008-0.343 (F_{ST}=0.002-0.113) (Appendix 4). Hierarchical AMOVA revealed a highly significant ($p<0.001$) proportion of genetic variance partitioned among rivers (4.12%), and within rivers (95.83%) (Table 3.2). A non-significant ($p>0.05$) component of variation was partitioned among temporal replicates of river collections (0.05%), and suggested temporally stable population structure among drainages

3.3.4 Relationships Among Rivers

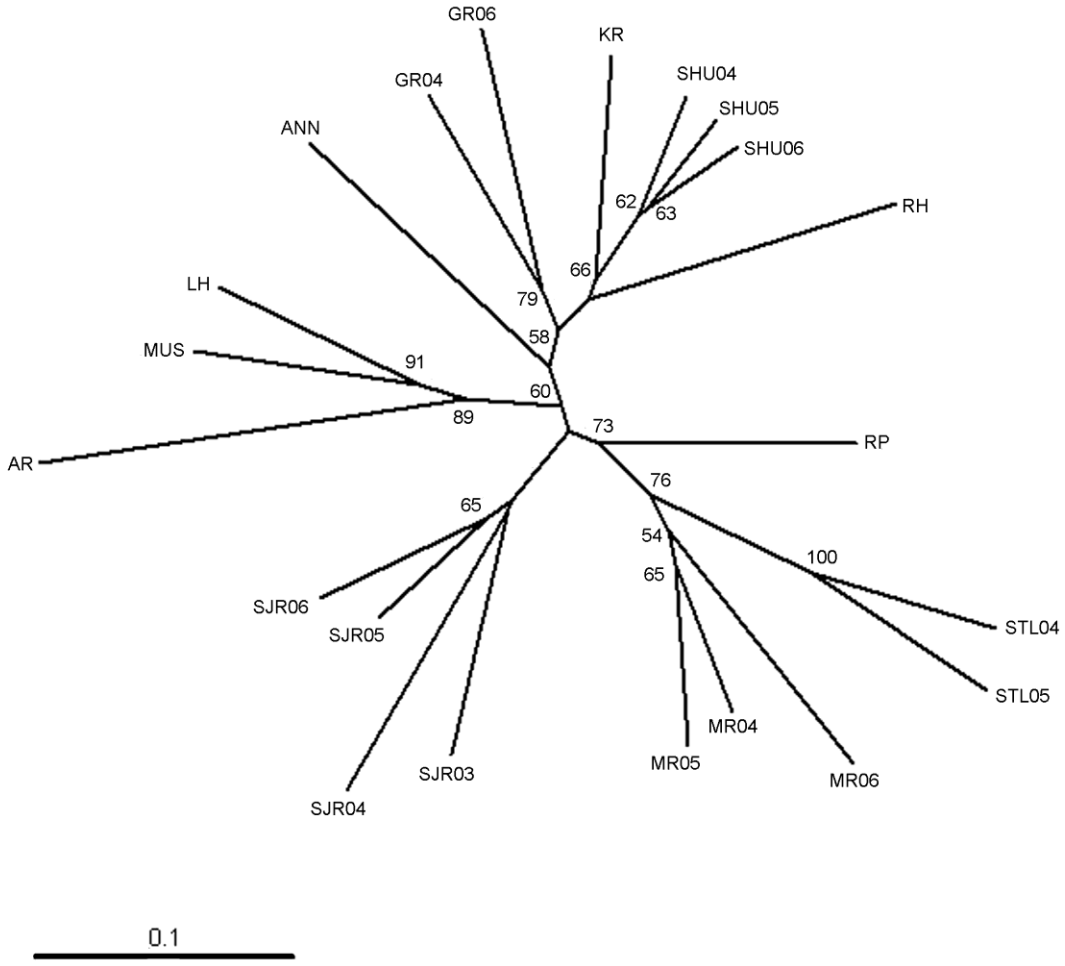
Neighbour joining analysis using D_C revealed that most drainages clustered by geographic region (Figure 3.2). Rivers tributary to the Gulf of St. Lawrence clustered together with 73% bootstrap support, while those on the Atlantic coast of Nova Scotia clustered together with 89% bootstrap support. Relationships among rivers tributary to

the Bay of Fundy were more complex. Shad from the Saint John River did not cluster with remaining Bay of Fundy rivers, but were positioned between the Gulf of St. Lawrence and Atlantic coast of Nova Scotia clusters (Figure. 3.2), with weak bootstrap support (<50%). The Shubenacadie and Kennetcook rivers clustered together with 66% bootstrap support, and suggested moderately close genetic affinities among these rivers, compared to the weak bootstrap support (<50%) observed among other drainages within the Bay of Fundy (i.e. Annapolis and Gaspereau rivers; Kennetcook and River Hebert).

3.3.5 Population Structure

Both STRUCTURE and BAPS identified seven clusters as the most likely number of genetically distinguishable groups among the 12 rivers sampled, suggesting evidence of metapopulation structure (i.e. genetically discrete local populations connected by migration; Hanski 1999). For BAPS, support for seven clusters was highly significant ($p<0.001$). Using STRUCTURE, the maximum value of $\ln Pr(X|K)$ was observed at $K=7$ (Figure 3.3a). The seven clusters identified by BAPS and STRUCTURE were identical in composition and corresponded to the following regions: i) St. Lawrence River, ii) Gulf of St. Lawrence (Miramichi River and River Phillip), iii) Atlantic coast of Nova Scotia (Musquodoboit, LaHave, and Annis rivers), iv) Annapolis River, v) Gaspereau River and River Hebert, vi) Shubenacadie and Kennetcook rivers, and vii) Saint John River (Figure 3.3b). Hierarchical STRUCTURE analyses (Vaha et al. 2007) conducted within each of the Gulf of St. Lawrence, Atlantic coast of Nova Scotia, Gaspereau-River Hebert, and Shubenacadie-Kennetcook clusters failed to detect further population structure, as the maximum value of $\ln Pr(X|K)$ within each cluster was observed at $K=1$ (data not shown). A highly significant (AMOVA; $p<0.001$) component of variation was detected among

Figure 3.2 Unrooted neighbor-joining tree for 21 collections of American shad from Atlantic Canada based on Cavalli-Sforza and Edwards (1967) chord distance (D_C) calculated using 13 microsatellite loci. Data were bootstrapped over loci with 10,000 replicates; bootstrap support values $\geq$50% are shown.

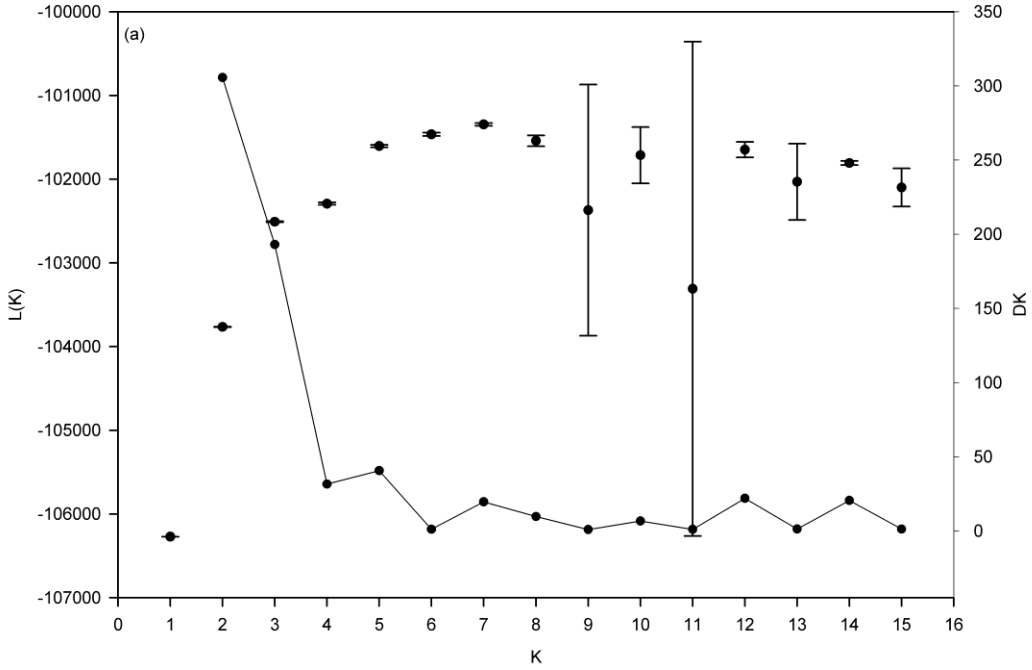

GR06
KR
SHU04
GR04
SHU05
ANN
SHU06
62
63
RH
66
LH
79
58
MUS
91
60
89
73
RP
AR
76
65
54
100
65
SJR06
STL04
SJR05
STL05
MR04
MR05
SJR03
MR06
SJR04
0.1

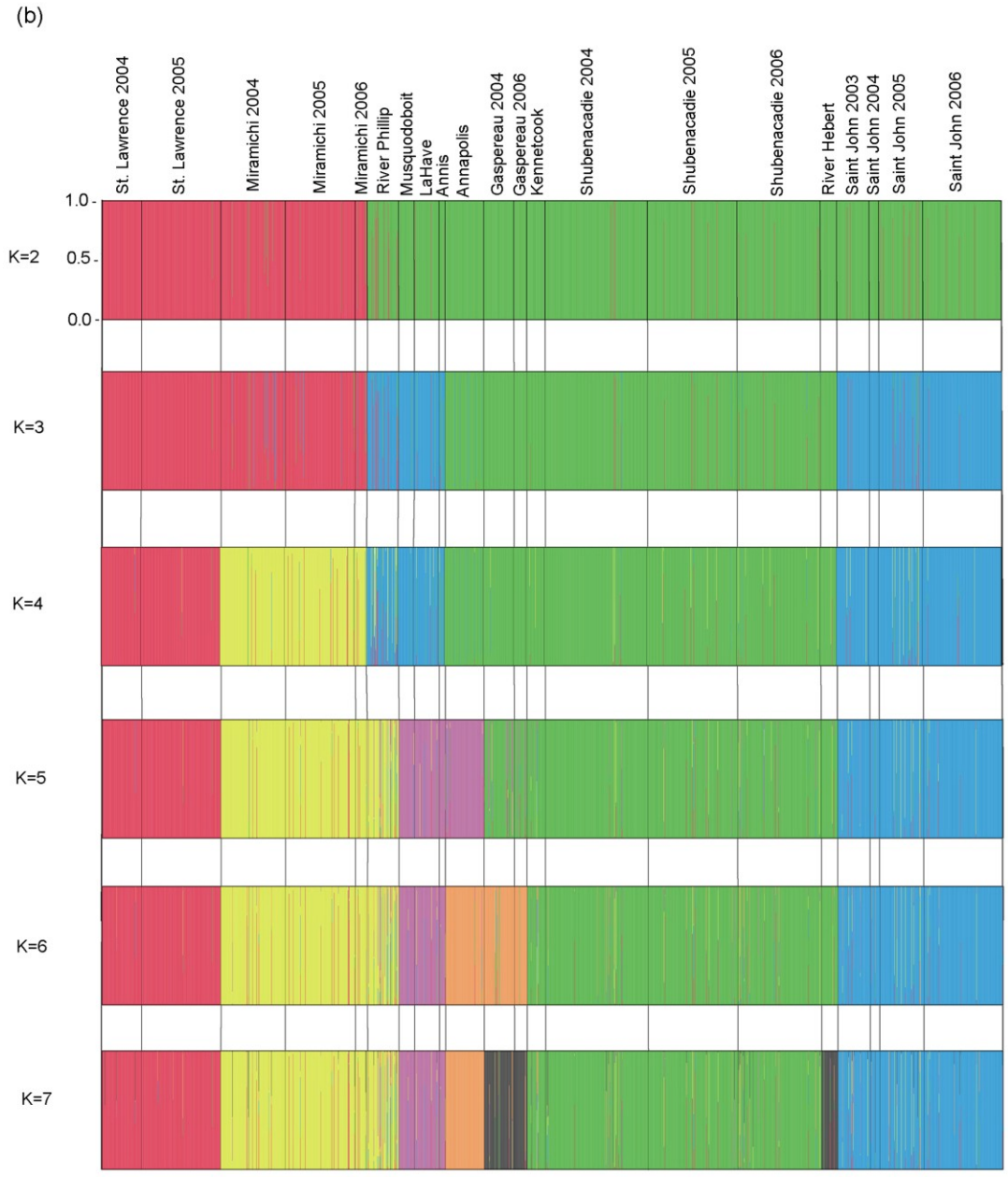
clusters (3.94%), and approximated that revealed among rivers when not grouped into clusters (4.12%; Table 3.2). The proportion of genetic variation partitioned among rivers within clusters (0.32%) was also highly significant ($p<0.001$), but considerably less than that observed among clusters. This result was not surprising given the significant genic differentiation detected among rivers.

Estimates of ΔK revealed the largest increase in the likelihood of the number of clusters at K=2 (Figure 3.3a), and suggested that the St. Lawrence and Miramichi rivers comprised one cluster, with River Phillip joining with all Atlantic coast of Nova Scotia and Bay of Fundy rivers in the second cluster (Figure 3.3b). However, hierarchical STRUCTURE analysis conducted independently for each of these clusters revealed the same seven clusters described above (data not shown).

BAPS admixture analysis revealed relatively few individuals (n=91; 3.7%) with significantly ($p<0.05$) admixed ancestry. Significantly admixed ancestry was most common in the Gaspereau/Hebert cluster (5.52%), followed by the Gulf of St. Lawrence (5.33%), Kennetcook/Shubenacadie cluster (4.10%), Atlantic coast of Nova Scotia (3.91%), Saint John River (2.88%), St. Lawrence River (1.54%) and Annapolis River (0%). Within rivers, the greatest proportion of admixed ancestry (~22%) was observed for the River Phillip collection. River Phillip admixture was largely due to inferred contributions of genomes from Atlantic Nova Scotia (8.6%) and Saint John River (5.5%) origins. River Phillip exhibited close affinities with both of these clusters, which remained undifferentiated in Bayesian model based clustering through K=4 (Figure 3.3b). Not until these clusters were separated at K=5 did River Phillip group with the Gulf of St. Lawrence cluster (Figure 3.3b). Remaining drainages demonstrated comparatively little

Figure 3.3 (a) Bayesian inference of number of American shad clusters (K) among 12 Canadian rivers using plateau of log probability of data L(K) (●; Pritchard et al. 2000), and ΔK (★; Evanno et al. 2005); (b) estimated population structure inferred from admixture analysis for K=2-7 clusters identified in Bayesian analyses. Individuals are represented by a thin vertical line which is partitioned into K-coloured segments representing an individual's estimated membership fractions from each of the identified clusters. Black lines separate individuals from different collections (labeled above).


(a)
L(K)
DK
K
-100000
-101000
-102000
-103000
-104000
-105000
-106000
-107000
350
300
250
200
150
100
50
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


(b)
St. Lawrence 2004
St. Lawrence 2005
Miramichi 2004
Miramichi 2005
Miramichi 2006
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau 2004
Gaspereau 2006
Kennetcook
Shubenacadie 2004
Shubenacadie 2005
Shubenacadie 2006
River Hebert
Saint John 2003
Saint John 2004
Saint John 2005
Saint John 2006
1.0
0.5
0.0
K=2
K=3
K=4
K=5
K=6
K=7

admixture from clusters other than the ones to which they were assigned (data not shown).

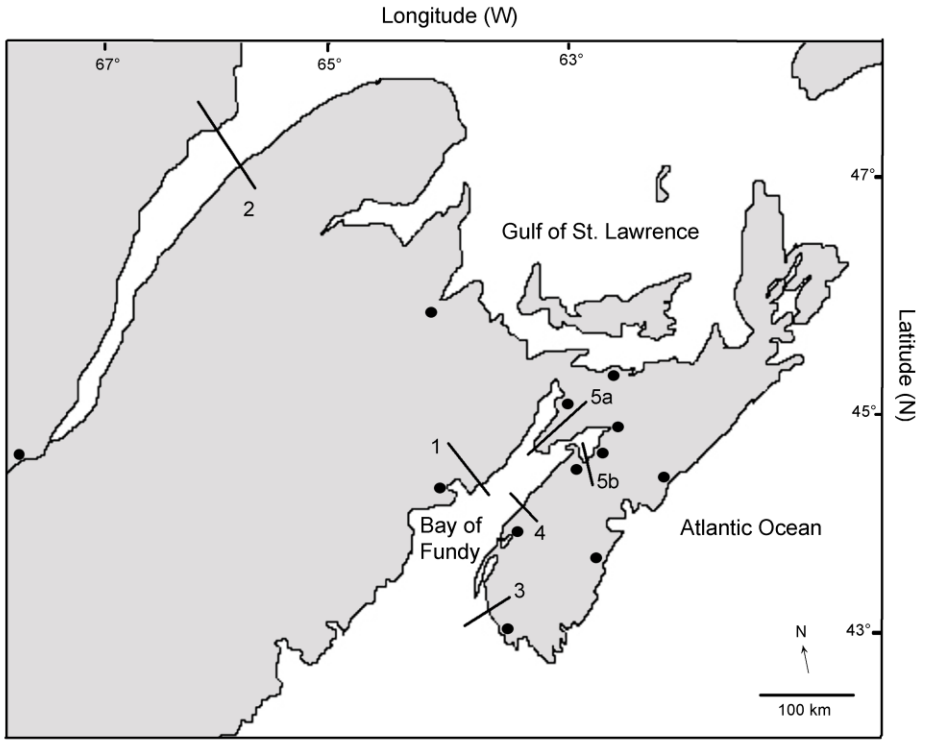
3.3.6 BARRIER ANALYSIS

BARRIER analysis revealed five barriers, four of which (1, 3, 4, 5) were located within the Bay of Fundy (Figure 3.4). The first (12 loci) and second (10 loci) consensus barriers isolated Saint John and St. Lawrence rivers, respectively (Figure 3.4). The third consensus barrier (11 loci) separated Annis River from the Annapolis, and effectively isolated Bay of Fundy rivers from those along the Atlantic coast of Nova Scotia and Gulf of St. Lawrence (Figure 3.4). The fourth consensus barrier (7 loci) isolated the Annapolis from the remaining Bay of Fundy rivers. Different configurations of connectivity among Bay of Fundy drainages (i.e. counterclockwise migration vs. direct route migration) did not shift position of these four consensus barriers. The fifth consensus barrier (7 loci) separated Shubenacadie River and River Hebert when connectivity among Bay of Fundy drainages was constrained to reflect a hypothesized counterclockwise migration route (i.e. Figure 3.4, barrier 5a) (Dadswell et al. 1987), but it shifted and isolated Gaspereau River from other Bay of Fundy drainages when this constraint was relaxed (i.e. Figure 3.4, barrier 5b).

3.3.7 ISOLATION BY DISTANCE

Mantel tests revealed a significant pattern of IBD among all rivers (Figure 3.5a), irrespective of the migration scheme used to assess distances among rivers within the Bay of Fundy (counterclockwise: $p=0.001$, $r=0.67$; direct: $p=0.003$, $r=0.66$). Among non Bay of Fundy rivers, IBD was highly significant and linear ($p<0.0001$; $r=0.98$; Figure 3.5b).

Figure 3.4 Major discontinuities in gene flow suggested by BARRIER using consensus barriers (≥7 loci), with order of importance given in Arabic numerals. Barrier #5 switched position depending on whether migration within the Bay of Fundy was constrained to a counterclockwise configuration (5a), or full connectivity (5b).


Longitude (W)
67°
65°
63°
47°
45°
43°
Latitude (N)
Gulf of St. Lawrence
Atlantic Ocean
Bay of Fundy
1
2
3
4
5a
5b
N
100 km

Within the Bay of Fundy, significant IBD was observed using counterclockwise distances ($p=0.02$; $r=0.59$; Figure 3.5b). When one notable outlier (i.e. the RH-SJR comparison) was removed from analysis, the IBD relationship strengthened considerably ($p<0.001$; $r=0.87$). Direct route distances among Bay of Fundy rivers resulted in non-significant IBD, and explained comparatively little of the variation in genetic differentiation among rivers ($r=0.14$).

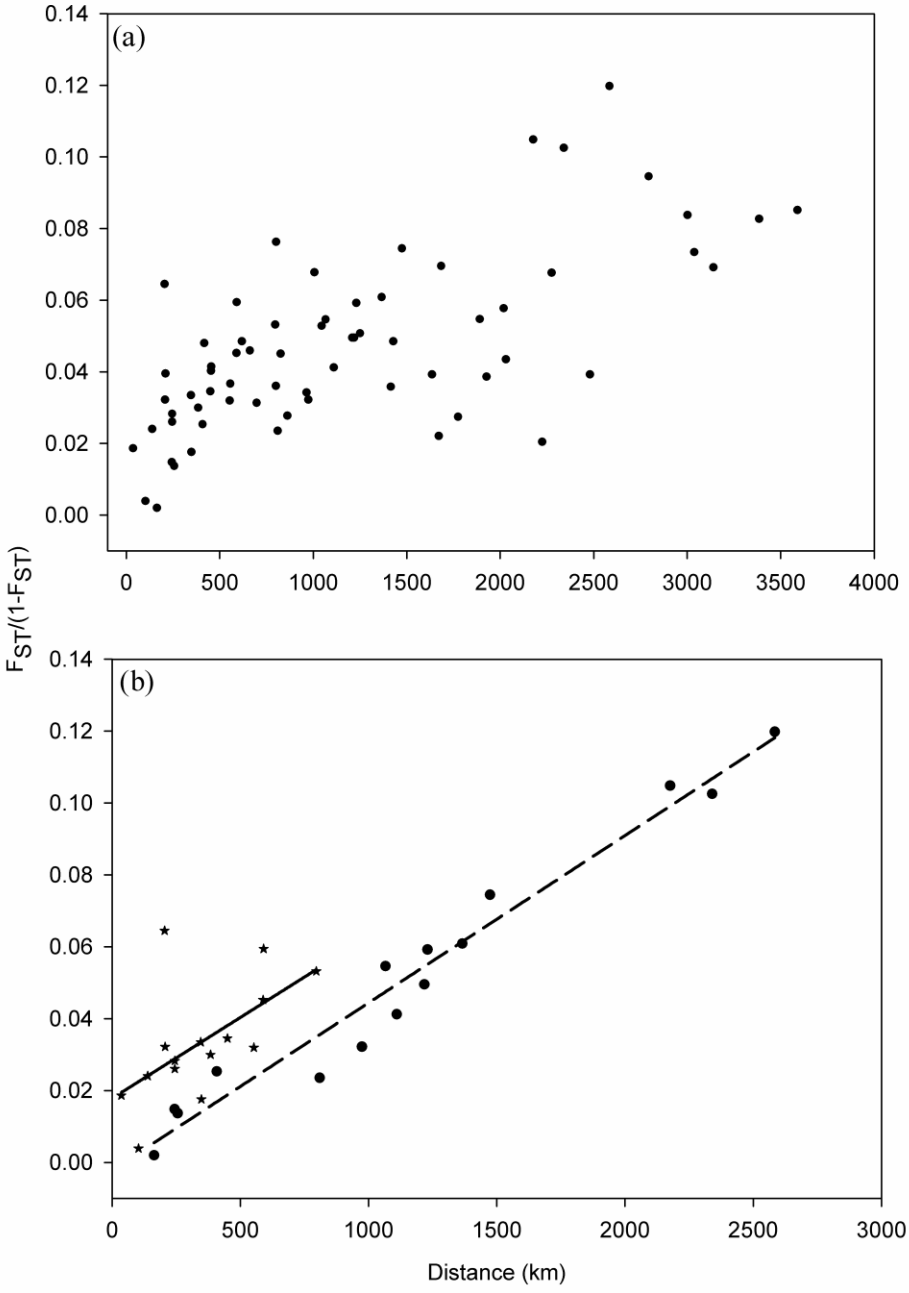
To determine whether distinct regional patterns of IBD exist, an analysis of covariance (ANCOVA; SYSTAT 10; SPSS 2000) was conducted among Bay of Fundy rivers (counterclockwise) and non Bay of Fundy rivers using geographic distance as a covariate. This permitted comparison of the degree of differentiation among regions while controlling for variation in genetic differentiation due to variation in geographic distance. Homogeneity of slopes was revealed among regions (non-significant region by covariate interaction; $p=0.92$), and suggested that regional patterns of IBD were similar. However, the degree of differentiation among rivers within the Bay of Fundy was greater than that among non Bay of Fundy rivers, regardless of the spatial scale of comparison (Figure 3.5b).

3.4 DISCUSSION

3.4.1 SPATIAL SCALE OF POPULATION STRUCTURE

Understanding the spatial scale of population structure can aid in effective management by serving as a precursor to defining conservation units. This survey of neutral genetic variation in American shad provides greater resolution than previous mtDNA based studies (Bentzen et al. 1989; Epifanio et al. 1995), and supports the

Figure 3.5 Relationship between pairwise genetic differentiation and geographic distance; (a) pairwise comparisons among all rivers using counterclockwise based distances among Bay of Fundy rivers, (b) linear regression of genetic differentiation on geographic distance among non Bay of Fundy (●), and Bay of Fundy rivers (counterclockwise distances) (★) revealed homogeneity of slopes (ANCOVA), but a higher degree of genetic differentiation for a given spatial scale among rivers within the Bay of Fundy.


(a)
(b)
$F_{ST}/(1-F_{ST})$
Distance (km)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
500
1000
1500
2000
2500
3000
3500
4000

hypothesis that individual rivers within the Canadian portion of the species range comprise genetically distinguishable spawning populations whose genetic compositions are temporally stable. I detected highly significant genic heterogeneity and significant population structure among all rivers, characterized by substantial levels of genetic differentiation and a significant pattern of IBD. Previous surveys of shad microsatellite variation revealed significant, but weak population structure among U.S. spawning runs (Brown et al. 2000), and failed to detect significant IBD (Waters et al. 2000). However, those studies examined only five loci and three spawning runs within close (<500 km) geographic proximity. This study incorporates data from 13 polymorphic loci, and 12 drainages from across a broad (>3500 km) geographic range, and confers sufficient statistical power to detect population structure among rivers at even weak levels of differentiation.

The comparatively low levels of genetic diversity observed for the St. Lawrence River is typical for populations at the periphery of species ranges (Eckert et al. 2008), and is consistent with a stepping stone model of population structure. The comparatively high levels of genetic diversity observed for the Musquodoboit River may reflect gene flow among rivers tributary to the Atlantic coast of Nova Scotia (see below), and is consistent with the lesser degree of genic differentiation observed between the LaHave and Musquodoboit Rivers.

Consistent with previous microsatellite (Brown et al. 2000) and mtDNA (Epifanio et al. 1995; Brown et al. 1996) based studies, temporal variation was not a significant source of genetic variation among rivers, and indicated that population structure was stable across at least short temporal scales (i.e. 2-3 years). This temporal stability may be

indicative of relatively large population sizes. Although accurate census population size estimates are not available, shad spawning runs within major Canadian rivers (e.g. Annapolis, Miramichi, Saint John, Shubenacadie) are expected to be on the order of 100s of thousands of fish (Chaput and Bradford 2003). Additionally, annual variation in allele frequencies (due to drift) are expected to be diminished for populations where spawners consist of several broadly overlapping year classes (Waples 1990; Waples and Teel 1990). This suggests that the temporal stability of shad population structure may be maintained through a combination of large population sizes, philopatry and spawning site fidelity (>97%; Melvin et al. 1986), variation in age at maturity (males: 3-4 yrs, females: 4-5 yrs; Walburg and Nichols 1967), and degree of iteroparity (>60%; Leggett and Carscadden 1978).

I did not observe evidence for temporally stable population sub-structuring among tributaries within the St. Lawrence, Miramichi, Shubenacadie, or Saint John River systems. Although I observed some evidence of genetic heterogeneity among tributaries within years, variation in allele frequencies across years within tributaries was also significant (and comparable in magnitude to differentiation observed among tributaries within years), and hierarchical AMOVAs detected a non-significant proportion of genetic variation partitioned among tributaries within these drainages. Further, Bayesian clustering analyses failed to detect sub-population structure, and revealed a single group within each of these drainages. My results are consistent with previous mtDNA based studies (Epifanio et al. 1995; Brown et al. 1996) which failed to observe population sub-structuring within U.S. drainages. Previous studies of morphological variation (Carscadden and Leggett 1975) and mark-recapture of hatchery reared specimens

(Hendricks et al. 2002) have suggested that shad home to their natal tributary. However, geochemical signatures in otoliths suggest that shad do not discriminate among tributaries within natal rivers when selecting spawning habitat (Walther et al. 2008). My results generally agree with this latter interpretation, and suggest that shad homing and spawning site fidelity is less effective at spatial scales finer than the drainage level.

3.4.2 METAPOPULATION DYNAMICS

Bayesian based estimates of seven genetically distinguishable shad population groups within the Canadian portion of the species range are likely conservative, as Bayesian methods may exhibit limited power under scenarios involving low levels of genetic differentiation (Faubet et al. 2007). For example, the Musquodoboit, LaHave, and Annis rivers exhibited significant genic heterogeneity, but were only weakly differentiated (F_{ST}=0.002-0.025), and were grouped into an 'Atlantic coast of Nova Scotia' cluster. Although the ΔK criterion supported the existence of only two genetic clusters within the Canadian portion of the species range, agreement between $\ln Pr(X|K)$ and ΔK based estimates have only been obtained using simulated data under simplistic conditions (i.e. island model of migration among four populations; Waples and Gaggiotti 2006). My data accord with previous empirical studies that revealed non-congruence between these methods (Piñeiro et al. 2007; Vaha et al. 2007). The inherent complexity of my data may explain the disparity observed between these estimators, and suggests that the ΔK statistic may have only detected 'deep-rooted' structure among populations (Piñeiro et al. 2007).

Nevertheless, Bayesian based clustering methods may have revealed an intriguing aspect of shad population dynamics which eluded previous investigations, but which is

important for shad conservation. Cumulatively, the clustering of neighboring (often geographically proximal) rivers exhibiting modest differentiation, low incidence of significant admixed ancestry among clusters, and geographic discontinuities in gene flow among regions suggest evidence for shad metapopulation structure (sensu Hanski 1999) on regional spatial scales (<500 km). The incidence of presumed metapopulations may be linked with habitat availability. While an assessment of habitat size/accessibility is beyond the scope of this study, shad from comparatively small rivers (relative to St. Lawrence or Saint John rivers) within regions clustered together. Although speculative, small rivers within regions may cumulatively support temporally stable spawning metapopulations, characterized by greater gene flow within than among regions. This scenario is consistent with my findings for rivers comprising the Atlantic coast of Nova Scotia and the Gaspereau-River Hebert clusters.

Shad metapopulation structure may also persist where comparatively small rivers are in close geographic proximity to larger spawning populations. This scenario is consistent with my findings for rivers comprising the Kennetcook-Shubenacadie cluster, and the Gulf of St. Lawrence cluster. Although significant allele frequency differences were detected among the Kennetcook and Shubenacadie Rivers, these rivers were only weakly differentiated. While a similar pattern was observed among rivers within the Gulf of St. Lawrence, this region presents additional complexity. Despite significant allele frequency differences, my results suggest that River Phillip is only weakly differentiated from the Miramichi River and rivers along Atlantic coast of Nova Scotia, perhaps through historical and/or contemporary gene flow. While a sixth consensus barrier was revealed between Gulf of St. Lawrence and Atlantic coast of Nova Scotia, it was weakly

supported (i.e. 6 loci), and suggested a greater degree of gene flow among these clusters than among others (data not shown). This result is consistent with the proportion of significant admixed ancestry from the Atlantic coast of Nova Scotia cluster detected within River Phillip.

3.4.3 Inference on Historical Biogeography

Observed patterns of shad population structure may also reflect aspects of the species historical biogeography in Atlantic Canada. Several analyses suggested that shad from the Saint John River exhibited closer affinities with those from the Gulf of St. Lawrence than other Bay of Fundy rivers. This pattern was evidenced by a neighbour-joining phenogram, but was also observed when a reduced number of clusters were specified during Bayesian analysis (K=2-4), and was consistent with the proportion of the River Phillip collection with significant admixed ancestry from the Saint John River. The Tantramar Marsh (spanning the border between New Brunswick and Nova Scotia) is a region of low elevation and topographic relief (Curry 2007), that may have been submerged post-glacially (glacial lake Acadia; Seaman 2004), providing a colonization route between the Bay of Fundy and Gulf of St. Lawrence. Although this scenario is speculative, an examination of historical and contemporary gene flow patterns between the Saint John and Gulf of St. Lawrence rivers would aid the interpretation of the genetic affinities observed between these drainages.

3.4.4 Isolation by Distance

I observed a significant increase in genetic differentiation with geographic distance among all Canadian rivers, consistent with a stepping stone model and the

geographic scale of population structure revealed herein. A strong linear IBD relationship (r=0.98) was observed among non Bay of Fundy rivers on spatial scales >2500 km, and within the Bay of Fundy using counterclockwise distances among rivers. Counterclockwise distances explained a significant and greater proportion of genetic variation among Bay of Fundy rivers than direct-route based distances, suggesting that pre-spawning adults follow a counterclockwise migration route within the Bay of Fundy. These results are consistent with previous findings based on the chronology of tag return data (Dadswell et al. 1983; 1987) and date of capture among Bay of Fundy rivers (Melvin et al. 1986), which suggested that non-spawning and post-spawning American shad in the Bay of Fundy follow a counterclockwise migration route. It has been suggested that shad migration within large semi-enclosed coastal regions follows the direction of the residual currents and coastline (Dadswell et al. 1987). Counterclockwise migration is consistent with the residual current flow of the Bay of Fundy (Bumpus and Lauzier 1965), and a similar pattern has been observed for shad migration within the Gulf of St. Lawrence (Dadswell et al. 1987). Considering that shad from rivers across their native range enter the Bay of Fundy during their annual feeding migration (Melvin et al. 1992), a counterclockwise migration pattern may allow individuals to take advantage of tidal stream transport (Weihs 1978), and the seasonal distribution of zooplankton within the Bay of Fundy (Locke and Corey 1989) to offset the energetic costs of extensive annual migrations.

While patterns of IBD were not significantly different among regions, a greater degree of differentiation was observed among Bay of Fundy rivers regardless of the geographic scale considered. This result is consistent with the greater number of

discontinuities in gene flow detected within the Bay of Fundy. Although the extent to which counterclockwise migration influences gene flow among Bay of Fundy rivers is uncertain, elevated levels of differentiation are probably recent in origin. Post-glacially (13,000-8600 ybp), the Bay of Fundy was a comparatively low energy (non-tidal), open marine system (Amos 1978); the dominant hydrodynamic features of the Bay of Fundy being geologically recent in origin (~6300 ±1100 ybp; Amos 1978). Given the average generation length for Canadian shad populations (5 years; Leggett and Carscadden 1978), and the potential role that residual currents may play, observed patterns of population structure within the Bay of Fundy may have only persisted for the past ~1260 (±220) generations.

3.4.5 MANAGEMENT IMPLICATIONS

This study suggests that effective management of Canadian shad populations requires a 'river-level' approach, and justifies current federal management practices. However, the detection of patterns consistent with the presence of shad metapopulation structure suggests that fisheries managers need to be concerned with the loss of genetic diversity at a range of spatial scales, from individual rivers (i.e. St. Lawrence, Annapolis, Saint John) to regions (i.e. Gulf of St. Lawrence, Atlantic coast of Nova Scotia). The elevated level of stock structure observed within the Bay of Fundy may warrant particular consideration should development of tidal power proceed within the region.

These results are consistent with previous studies, which failed to detect sub-population structure within rivers (Epifanio et al. 1995; Brown et al. 1996). However, we cannot reject the possibility that the exclusion of shad from spawning habitat above Mactaquac Dam in the Saint John River resulted in the loss of a genetically

distinguishable sub-population. Adult spawners heading further upstream but encountering the dam may have spawned in various tributaries below Mactaquac Dam; resulting in introgression which may have yet to reach equilibrium. Considering the significant level of genic differentiation detected among all rivers examined, it is likely that previous extirpations of shad from St. Croix River, NB, and Peticodiac River, NB, resulted in the loss of genetically distinguishable populations. Estimation of the magnitude of the loss would require knowledge of whether or not these populations were members of a larger metapopulation, perhaps including the Saint John River. It will be of great interest to monitor the origin of migrant fish contributing to any future natural recolonizations of these rivers as access to spawning habitat is re-established.

Using fisheries data as an indicator, Leim (1924) noted the relative scarcity of shad spawning populations along the Atlantic coast of Nova Scotia, and attributed this to the absence of large accessible rivers required for successful reproduction. My results suggest that 'large' drainages are not an absolute requirement for either self-sustaining spawning activity, or the persistence of spawning metapopulations. This possibility warrants further consideration of historical accounts of shad from additional 'small' drainages (see Chaput and Bradford 2003) along the Atlantic coast of Nova Scotia (e.g. Mersey, Medway, and St. Mary's Rivers) and within the Bay of Fundy (e.g. Avon, St. Croix (Nova Scotia), LaPlanche, Nappan, and Maccan Rivers).

CHAPTER 4

RANGEWIDE POPULATION STRUCTURE IN AMERICAN SHAD: HISTORICAL CONTINGENCY OR ANTHROPOGENIC INFLUENCE?

4.1 INTRODUCTION

The long-term persistence of species is to a great extent contingent on the ability of populations to mount variable responses to perturbations, the breadths of which are largely dependent on the amount of heritable variation present at the population level (Allendorf and Lundquist 2003). Maintenance of genetic variation and connectivity among populations from across the range of habitats in which species persist can increase the capacity for evolutionary change (Mace and Purvis 2008). However, populations are not necessarily equivalent in their responses to future environmental conditions (Petit et al. 1998), and the prioritization of populations for focusing limited conservation resources is often desirable, but requires knowledge of the magnitude and spatial distribution of genetic variation from across a species' range.

For many species, the magnitude and spatial distribution of neutral genetic variation reflects the combined effects of historical demographic factors, contemporary microevolutionary processes, and anthropogenic influences. Historical demographic factors associated with climatic change have had a profound influence on the spatial distribution of intraspecific neutral genetic variation (Beaumont 1999; Storz and Beaumont 2002; Bos et al. 2008), constraining the amount of variation upon which microevolutionary forces may act. Recurrent glacial cycles during the Pleistocene epoch reduced intraspecific genetic variation through successive population bottleneck and

founder events (Hewitt 2000), while different modes of post-glacial dispersal (Ibrahim et al. 1996) influenced the initial spatial structure of populations during interglacial (10-12ky duration; Dawson 1992) range expansions. Despite the significant role of historical demography, meaningful interpretation of contemporary spatial patterns of neutral genetic variation requires consideration of the effects of microevolutionary processes and anthropogenic factors. Contemporary spatial patterns of neutral genetic variation may reflect the influences of gene flow and/or genetic drift since post-glacial colonization, but may be distorted by habitat fragmentation or translocations of individuals among reproductively isolated demes.

Shad provide an interesting model to investigate the influence of historical demography on contemporary spatial patterns of genetic variation, as reproductively isolated (Melvin et al. 1986; Hendricks et al. 2002) and genetically distinguishable spawning runs (Bentzen et al. 1989) are continuously distributed among major rivers from Florida to Quebec (Walburg and Nichols 1967), and the species' contemporary distribution (~30°N-50°N) is evenly divided among non-glaciated and formerly glaciated regions by the maximal extent of the Laurentide ice sheet during the Wisconsinan glaciation (~40°N; Schmidt 1986). Shad are believed to have been confined to the southeastern United States during the last glacial maximum (LGM; 23-18 ky bp) (Bentzen et al. 1989). Post-glacial dispersal from this region may have followed a stepping stone process (Ibrahim et al. 1996) as habitat may have become sequentially suitable for colonization, and may have manifested as successive reductions in genetic variation with increasing latitude. However, spatial patterns of shad neutral genetic variation cannot be interpreted solely in the context of historical demography and post-

glacial dispersal. Range-wide declines in shad abundance attributable to anthropogenic influences (i.e. overfishing, dams, pollution; Bilkovic et al. 2002a) have prompted over a century of supportive breeding and stocking activities (including stock transfers among U.S. rivers; Hendricks 2003). The ultimate impacts of historical bottleneck/founder events, contemporary gene flow/drift, and stocking practices on spatial patterns of neutral genetic variation in shad remain unknown. However, one might predict that historical demographic factors have resulted in lower levels of genetic variation among spawning populations in formerly glaciated areas, and postulate that stock transfers have reduced levels of genetic differentiation (i.e. genetic swamping; Bouzat et al. 2009) among some U.S. spawning runs.

Substantial resources have been allocated to shad restoration measures (e.g. hatchery supplementation, modification of fish passage, dam removal) with varying success (reviewed in Hendricks 2003; Cooke and Leach 2003; Olney et al. 2003; St. Pierre 2003; Weaver et al. 2003). Resolving the distribution of shad neutral genetic variation may benefit future restoration efforts, as knowledge of the relative contributions of populations to species genetic variation could aid prioritization of populations for management, leading to an effective conservation strategy, and the long-term persistence of shad.

Prior molecular studies of shad mainly employed mitochondrial DNA (mtDNA) to examine spatial patterns of neutral genetic variation (e.g. Bentzen et al. 1989; Epifanio et al. 1995; Waldman et al. 1996); those studies that have examined microsatellites were limited to few loci and few populations (Brown et al. 2000; Waters et al. 2000). In this chapter I examine the magnitude and spatial distribution of neutral genetic variation at 13

microsatellite loci among 33 shad spawning runs from across the species' range, and i) assess whether the distribution of neutral genetic variation across formerly glaciated habitats is consistent with stepwise post-glacial colonization, ii) assess evidence for different patterns of spatial genetic variation across the range of the species, and whether such patterns are consistent with the influence of historical demography or alternative management practices, and iii) examine the relative contribution of each population to overall species genetic variation.

4.2 MATERIALS AND METHODS

4.2.1 SAMPLE COLLECTIONS

To capture as much intraspecific genetic variation as possible, shad were collected from across their native range. Sampling from 2003-2006 employed a combination of trapnets, gillnets, seines, electrofishing, and angling, and provided 4575 fin clips from 33 rivers (e.g. spawning runs) (Table 4.1; Figure 4.1). Samples consisted of adults captured in rivers above the influx of saltwater, with the following exceptions: the Kennetcook River collection was comprised primarily of juvenile specimens, while River Phillip and Miramichi River (2004-2005) collections consisted of adults obtained from tidal freshwater. These collections were considered representative of adult spawning populations for reasons discussed in Chapter 3, and were retained for analyses. Fin clips were preserved in 95% ethanol until DNA extraction.

Table 4.1 Sampling locations and sample sizes (N) for American shad collected from 2004-2006 among 33 spawning runs from across the species' range.

River No.	Code	Spawning population	Latitude	2003	2004	2005	2006	Total
1	STL	St. Lawrence River, PQ	49°09'41 N		120	221		341
2	MR	Miramichi River, NB	47°08'31 N		212	196	33	441
3	RP	River Phillip, NS	45°47'34 N		16	22	51	89
4	MUS	Musquodoboit River, NS	44°42'12 N		12	20	13	45
5	LH	LaHave River, NS	44°17'34 N		12	15	41	68
6	AR	Annis River, NS	43°52'32 N		10		7	17
7	ANN	Annapolis River, NS	44°19'16 N		111			111

Table 4.1 *continued*

River No.	Code	Spawning population	Latitude	2003	2004	2005	2006	Total
8	GR	Gaspereau River, NS	45°05'22 N		85		35	120
9	KR	Kennetcook River, NS	45°02'11 N			50		50
10	SHU	Shubenacadie River, NS	45°15'37 N		284	246	230	760
11	RH	River Hebert, NS	44°43'37 N		23	22		45
12	SJR	Saint John River, NB	45°15'34 N	95	27	127	222	471
13	MER	Merrimack River, MA	42°49'23 N		100			100
14	CON	Connecticut River, CT	41°16'58 N		100			100
15	HUD	Hudson River, NY	40°41'45 N		100			100

Table 4.1 *continued*

River No.	Code	Spawning population	Latitude	2003	2004	2005	2006	Total
16	DEL	Delaware River, PA	39°06'44 N		25			25
17	NAN	Nanticoke River, MD	38°10'08 N			13		13
18	SUS	Susquehanna River, MD	38°27'44 N			59		59
19	PAT	Patuxent River, MD	38°15'29 N			63		63
20	POT	Potomac River, MD	38°00'15 N			135		135
21	RAPP	Rappahannock River, VA	37°29'36 N			95		95
22	YOR	York River, VA	37°12'33 N			101		101
23	JAM	James River, VA	36°58'07 N			99		99

Table 4.1 *continued*

River No.	Code	Spawning population	Latitude	2003	2004	2005	2006	Total
24	ROA	Roanoke River, NC	35°55'22 N			50		50
25	TAR	Tar River, NC	35°13'52 N			100		100
26	NEU	Neuse River, NC	35°04'35 N			50		50
27	CF	Cape Fear River, NC	33°55'43 N			100	100	200
28	WACC	Waccamaw River, SC	33°13'59 N			100		100
29	COOP	Cooper River, SC	32°48'28 N			65	28	93
30	ED	Edisto River, SC	32°31'55 N			99		99
31	SAV	Savannah River, GA	32°02'53 N			23	18	41

Table 4.1 *continued*

River No.	Code	Spawning population	Latitude	2003	2004	2005	2006	Total
32	ALT	Altamaha River, GA	31°18'53 N			100	100	200
33	STJ	St. John's River, FL	30°24'29 N			94	100	194

4.2.2 LABORATORY PROTOCOLS

DNA isolation followed the method of Elphinstone et al. (2003), modified for automation in 96 well format using a Perkin Elmer MPII liquid handler. Specimens were genotyped across 14 polymorphic microsatellite loci developed for *A. sapidissima* ((*AsaB020, AsaD029, AsaD042, AsaD429*; Julian and Bartron 2007) (*Asa16*; Brown et al. 2000) (*Asa2, Asa4, Asa8*; Waters et al. 2000)), *A. fallax* (*Af6, Af13, Af20*; Faria et al. 2004), *A. alosa* (*Aa14, Aa16*; Faria et al. 2004), and *A. pseudoharengus* (*Aps2A;* Bentzen and Paterson 2005). Details of polymerase chain reaction (PCR) conditions were as previously reported (Chapter 3). Alleles were scored visually by reference to a molecular weight size standard constructed from pUC18 derived PCR fragments. Gel images were scored with reference to positive controls, negative controls, and redundant samples to ensure consistency and reproducibility of genotypes across collections and loci.

4.2.3 DATA ANALYSIS

Data conformance to model assumptions

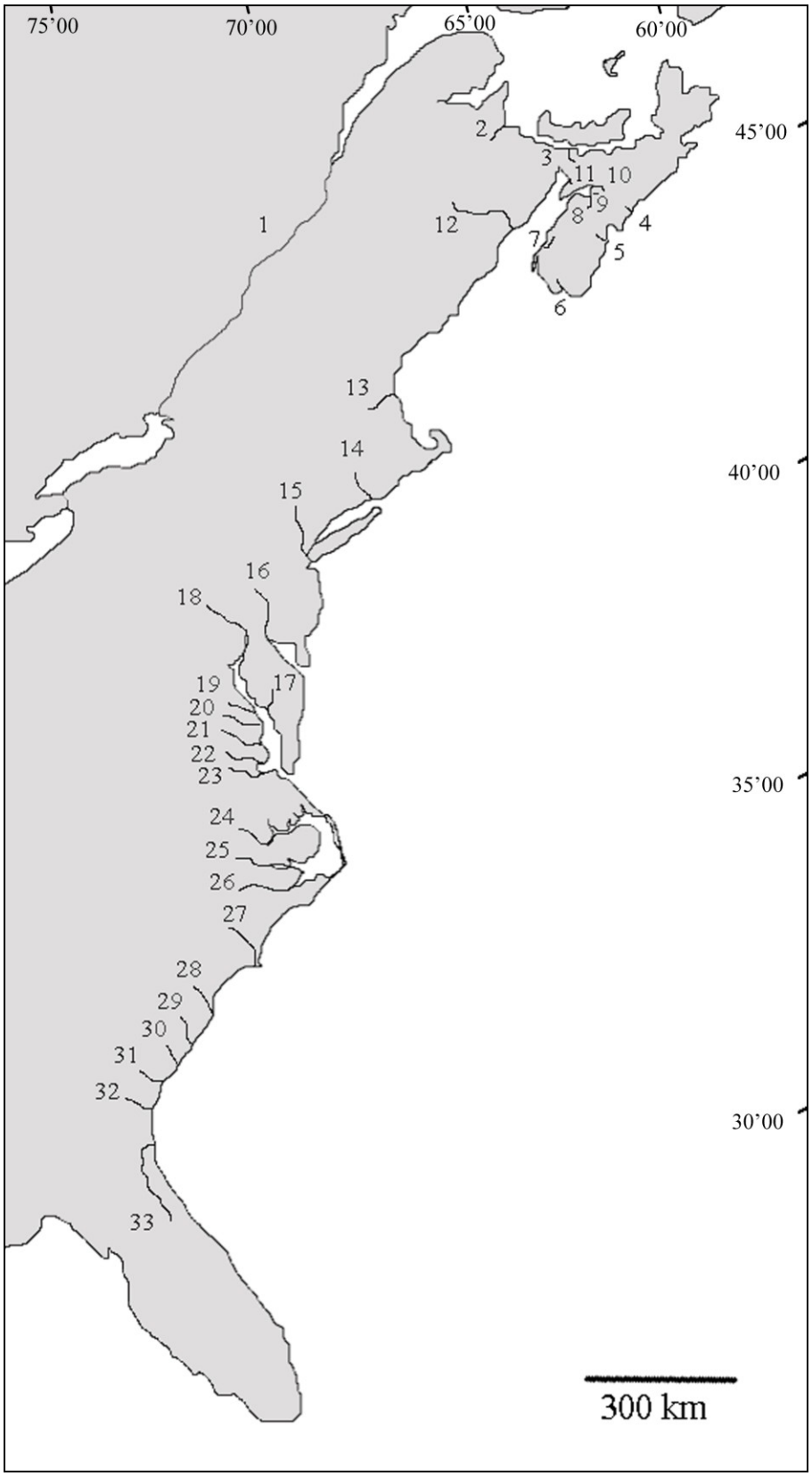
The presence of genotyping artefacts was evaluated using Microchecker v.2.2.3 (van Oosterhout et al. 2004). Tests for departures from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) were performed with GENEPOP v. 4.0.6 (Rousset 2007) using default parameters for all tests. Sequential Bonferroni adjustments were used to judge significance levels for all simultaneous tests (Holm 1979; Rice 1989). Selective neutrality of the microsatellite markers used in this study was evaluated by assessing relative variance in repeat number (lnRV) and heterozygosity (lnRH) (Schlötterer 2002; Schlötterer and Dieringer 2005).

Spatial trends in genetic diversity

For each river a multilocus estimate of observed heterozygosity (H_O), an unbiased estimate of heterozygosity (H_E; Nei 1978), and F_{IS} (Weir and Cockerham 1984) was calculated using the program GENETIX v. 4.05 (Belhkir et al. 2004). A multilocus estimate of allelic richness (R) was estimated for each river using Fstat 2.9.3.2 (Goudet 1995; 2001) standardized to a minimum sample size of 32 individuals (LeBerg 2002). Collections from the Annis, Delaware, and Nanticoke rivers were excluded from allelic richness estimation because of small sample sizes for those rivers (Table 4.1).

Multilocus estimates of H_O and R were examined for patterns of spatial variation using latitude of river mouth. Permutation tests (FSTAT; 10,000 permutations) were used to determine whether levels of genetic diversity differed among rivers from formerly

Figure 4.1 Map of Atlantic coast of North America displaying sampling locations for collections of American shad: 1-St.Lawrence River, QC; 2-Miramichi River, NB; 3-River Phillip, NS; 4-Musquoboboit River, NS; 5-LaHave River, NS; 6-Annis River, NS; 7-Annapolis River, NS; 8-Gaspereau River, NS; 9-Kennetcook River, NS; 10-Shubenacadie River, NS; 11-River Hebert, NS; 12-Saint John River, NB; 13-Merrimack River, MA; 14-Connecticut River, CT; 15-Hudson River, NY; 16-Delaware River, PA; 17-Nanticoke River, MD; 18-Susquehanna River, MD; 19-Patuxent River, MD; 20-Potomac River, MD; 21-Rappahannock River, VA; 22-York River, VA; 23-James River, VA; 24-Roanoke River, NC; 25-Tar River, NC; 26-Neuse River, NC; 27-Cape Fear River, NC; 28-Waccamaw River, SC; 29-Cooper River, SC; 30-Edisto River, SC; 31-Savannah River, GA; 32-Altamaha River, GA; 33-St. John's River, FL.


75’00
70’00
65’00
60’00
45’00
40’00
35’00
30’00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
300 km

glaciated and non-glaciated regions (i.e. above and below 40°N, respectively; Schmidt 1986). However, the *a priori* choice of this latitude may be inappropriate for the examination of spatial variation in genetic diversity among contemporary spawning runs. Sufficient time may have elapsed since post-glacial colonization of habitats within close proximity of the glacial margin (e.g. Hudson River) to restore initial losses of genetic diversity anticipated following population founding events. Therefore, I conducted piecewise linear regressions as a non-apriori method of assessing latitude associated trends in genetic diversity. Piecewise regressions (SYSTAT v.10; SPSS 2000) were performed iteratively (1° increments) across latitudes from 30°N to 49°N for each of H_O and R to resolve the 'latitudinal breakpoint' at which divergent patterns of genetic diversity become apparent. The latitude at which the greatest proportion of global variation in genetic diversity was explained by piecewise regression was taken as the latitudinal breakpoint.

Clinal variation in allele frequencies

Recent studies have demonstrated the influence of drift on genetic variation during range expansion, and have shown how low frequency alleles can 'surf' to high densities on the wave of advance and establish allele frequency clines (reviewed in Excoffier and Ray 2008). To explore evidence for allele frequency clinal variation, I estimated allele frequency distributions by river for each locus and conducted linear regressions of the most common allele at each locus against latitude using SYSTAT v.10 (SPSS 2000).

Genetic differentiation

The statistical power and realized α-error for assessing the null hypothesis of genetic homogeneity among rivers was assessed using POWSIM (Ryman and Palm 2006). Allelic heterogeneity among rivers was assessed via genic tests conducted with GENEPOP v.4.0.6 (Rousset 2007), and with default parameters for all tests. Tests were combined across loci or collections using Fisher's method. Hierarchical AMOVA was conducted among temporally replicated river collections to partition components of genetic variation among rivers, among years within rivers, and among individuals within rivers, using a permutation procedure (10,000 iterations) (Arlequin 3.1; Excoffier et al. 2005).

Genetic differentiation was estimated by computing overall and pairwise F_{ST} (θ; Weir and Cockerham 1984) using FSTAT 2.9.3.2(Goudet 1995; 2001). The effects of variation in genetic diversity on genetic differentiation (Hedrick 2005) were controlled for by calculating standardized estimates of differentiation (F'_{ST}) using RECODEDATA v. 0.1 (Meirmans 2006), which recoded the data so that there were no shared alleles among rivers. Recoded data were then used to estimate $F_{ST(max)}$ for each pairwise comparison using FSTAT (Goudet 1995; 2001). Standardized pairwise estimates of differentiation were then calculated as $F'_{ST} = F_{ST} / F_{ST(max)}$ following Hedrick (2005).

Relationships among rivers

Genetic affinities among rivers were examined using principal coordinates analysis (PCA) of the pairwise genetic distance matrix for D_A (Nei et al. 1983), and as implemented in GenAlEx v.6.0 (Peakall and Smouse 2006). This method provided a

perspective of the underlying structure of the genetic distance matrix without imposition of a bifurcating evolutionary history, and may more accurately represent the data than neighbour-joining trees when there is considerable genetic exchange between and among geographically proximal demes (Cavalli-Sforza et al. 1994), as is probably the case in this study. Although multivariate ordination techniques are useful for visualizations of population associations, evaluating the consistency and statistical significance of these relationships is problematic, and these methods are best suited to exploratory analyses rather than precise statistical inference (Pritchard et al. 2000). More efficient methods include maximum likelihood based assignment procedures (Paetkau et al. 1995; Rannala and Mountain 1997) and Bayesian clustering models (Pritchard et al. 2000).

Population structure

Two Bayesian model-based clustering methods, implemented in STRUCTURE v. 2.2 (Pritchard et al. 2000; Falush et al. 2003) and BAPS v. 5.1 (Corander et al. 2006), were used concurrently and in a hierarchical fashion to infer the number of genetically homogenous clusters, as advocated by Latch et al. (2006). For STRUCTURE analysis, a burn-in of 50,000 replicates was followed by 250,000 replicates of the Markov chain Monte Carlo (MCMC) simulation, employing the admixture model and correlated allele frequencies among populations. To assess convergence of the log probability of data Pr(X|K) among rivers, three iterations of this parameter set were performed for K (number of clusters) from 1-13, allowing an estimation of the most likely number of clusters. Both the plateau of likelihood values (Pritchard et al. 2000) and ΔK (i.e. second order rate of change between successive K values; Evanno et al. 2005) were estimated.

For the BAPS analysis, the mixture model was first applied to cluster groups of individuals based on their multilocus genotypes. Three iterations of K (1-13) were conducted among rivers to determine the optimal number of genetically homogeneous groups. Admixture analysis (Corander and Marttinen 2006) was then applied to estimate individual admixture proportions with regards to the most likely number of K clusters identified (Corander and Marttinen 2006), and were visualized using DISTRUCT v. 1.1 (Rosenberg 2004).

BARRIER analysis

BARRIER v. 2.2 (Manni et al. 2004) was used to define geographic discontinuities in gene flow among rivers. Barriers were first computed using a multilocus F_{ST} matrix along with geographic coordinates for each river sampled to represent genetic differentiation among spawning runs. F_{ST} matrices for each of the 13 loci were then used separately to identify 'consensus barriers' (i.e. $\geq$ 7 loci) to assess the robustness of the barriers identified using multilocus F_{ST}. Connectivity among rivers was largely restricted to reflect a stepping stone model of population structure. However, connectivity among Bay of Fundy rivers was constrained to reflect the counterclockwise migration route supported in Chapter 3, while full connectivity was permitted among Chesapeake Bay drainages following the most 'direct route' distance and incorporating trans-Chesapeake Bay measures.

Isolation by distance

An analysis of isolation by distance (IBD) was conducted among rivers to test for correlations between geographic distance and genetic differentiation using 10,000

permutations of the Mantel test as implemented in the web-service program IBDWS v. 3.15 (http://www.ibdws.sdsu.edu) (Jensen et al. 2005). Pairwise F_{ST} values were linearized as ($F_{ST}/(1-F_{ST})$) following Rousset (1997). Geographic distance between rivers was measured as the most direct route following within 5 km of the coastline using Google Earth 2006, as in Chapter 3. This distance was chosen as pre-spawning adults are known to migrate within a few kilometers of the coastline during their spawning migration (Leggett and Whitney 1972; Neves and Depres, 1979). Distances between Bay of Fundy rivers were measured following the counter-clockwise migration route supported in Chapter 3, while distances among Chesapeake Bay drainages were measured using the most direct route between rivers.

The distribution of residuals generated from a linear regression of the IBD pattern among U.S. spawning runs was investigated to determine whether stock transfers may have resulted in a lesser degree of genetic differentiation between donor and recipient spawning stocks than what might be predicted based on geographic distance alone. Historical shad stocking records (e.g. annual agency completion reports; U.S. Fisheries Commission internal documents, etc.) from 1871 to the present, and summary reports (Leach 1925; Robbins and Watts 1981; Hendricks 2003) were reviewed to understand which pairs of spawning populations have been the subject of stock transfers (i.e. donor and recipient spawning populations) during the course of shad stocking activities.

To determine whether different patterns of population structure were evident across the species' range, an analysis of covariance (ANCOVA; SYSTAT 11; SPSS 1997) was conducted among life history types (semelparity vs. iteroparity), glacial histories (formerly glaciated vs. non-glaciated rivers), and management practices

(stocked vs. non-stocked) using standardized geographic distances as a covariate. This permitted comparison of the degree of differentiation among life history types, glacial histories, and management practices while controlling for variation in genetic differentiation due to variation in geographic distance.

Contributions to genetic diversity

Because individual populations may contribute disproportionately to measures of total genetic diversity (Petit et al. 1998), the amount and proportion of genetic diversity within and among all spawning runs was estimated using the program CONTRIBUTE (Petit et al. 1998). This program estimates the relative contribution of each population (k) to total gene diversity (Nei 1973) across all populations surveyed (n) by comparing the total diversity of all populations to the diversity excluding the k^{th} population (Page et al. 2004). The relative contribution of the k^{th} population to total H_O (C_t) was partitioned into estimates of H_O within the k^{th} population (C_s; diversity), and the genetic divergence of the k^{th} population from other populations (C_d; differentiation). The contribution of each population to total R (C_{rt}) was assessed as a relative measure of the number of alleles observed (C_{rs}; diversity), and according to whether the population possessed alleles not present in other populations (C_{rd}; divergence) (Page et al. 2004). The Annis, Delaware, and Nanticoke rivers were excluded from these analyses due to sample size restrictions (Table 4.1).

4.3 RESULTS

4.3.1 DATA CONFORMANCE TO MODEL ASSUMPTIONS

Although within year collections for several rivers failed to meet target sample sizes ($n \geq 50$) (Table 4.1), analysis of molecular variance (AMOVA) revealed non-significant ($p>0.05$) genetic variation among years within all rivers sampled (see below). Temporally replicated collections within all rivers were consequently pooled prior to analyses. Locus *Aa16* was removed from analyses because Microchecker detected null alleles for >50% of collections examined. Remaining loci were retained for analyses, as evidence for null alleles was sporadically distributed among loci and collections. Exact tests revealed that genotypic frequencies were largely in accordance with HWE ($p>0.05$; sequential Bonferroni correction for 33 comparisons). However, departures from HWE remained for 46 locus-collection comparisons after sequential Bonferroni adjustment, 41 of which were due to heterozygote deficiencies resulting from either sporadic presence of null alleles, or from potential Wahlund effects associated with the collection of shad from multiple sites within rivers (i.e. St. Lawrence, Miramichi, Shubenacadie, Saint John River systems) (see Chapter 3). Exact tests of LD revealed that loci were physically unlinked and statistically independent ($p>0.05$; sequential Bonferroni correction for 2574 comparisons). Relative variance in repeat number (lnRV) and heterozygosity (lnRH) failed to detect outlier loci (i.e. markers were normally distributed; data not shown), and provided no evidence of non-neutrality among the loci examined.

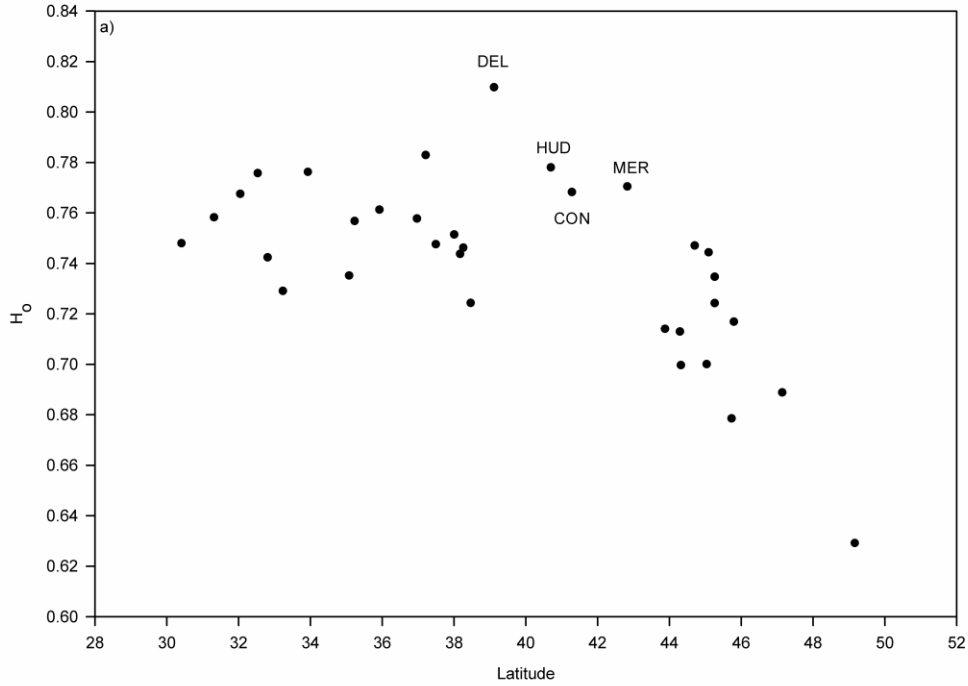
4.3.2 SPATIAL TRENDS IN GENETIC DIVERSITY

The degree of microsatellite polymorphism among the 33 rivers varied greatly depending on the locus and spawning run considered. The number of alleles per locus ranged from 10 (*Af20*) to 43 (*Aa14*), with 11 loci exhibiting ≥15 alleles; Appendix 5). Observed heterozygosity varied between 0.63 (St. Lawrence River) and 0.81 (Delaware River), while allelic richness ranged from 7.07 (St. Lawrence River) to 11.69 (Edisto River) (Appendix 5). Private alleles were observed among 14 spawning runs for one or more loci, but were always rare in frequency (<0.05; data not shown). Allele frequency distributions across the species' range for the 13 loci examined in this study are provided in Appendix 6.

Across the species range, broadly similar patterns of reduced genetic diversity with increasing latitude were observed for both H_O and R (Figure 4.2). Permutation tests revealed significantly ($p<0.05$) greater levels of genetic diversity among spawning runs from the non-glaciated portion of the species range than those from formerly glaciated regions. Spawning runs south of the glacial maximum (40°N) exhibited a significantly ($p<0.005$) greater average H_O (0.76) than those to the north (0.71), and a highly significant ($p<0.001$) greater average R (10.90) than those from formerly glaciated regions (8.58).

Although genetic diversity was generally lower among spawning runs within the formerly glaciated portion of the species range, values of H_O (Figure 4.2a) and R (Figure 4.2b) for the Hudson, Connecticut, and Merrimack Rivers were similar to those for spawning runs from non-glaciated regions. This generated uncertainty about where to place the boundary between the two different latitudinal trends observed across the

Figure 4.2 Geographic distribution for multilocus estimates of genetic diversity across the species' range of American shad; a) observed heterozygosity (H_O), b) allelic richness (R). The position of the Hudson, Connecticut, and Merrimack collections are labeled (see text).


a)
DEL
HUD
MER
CON
H_o
Latitude
0.84
0.82
0.80
0.78
0.76
0.74
0.72
0.70
0.68
0.66
0.64
0.62
0.60
28
30
32
34
36
38
40
42
44
46
48
50
52

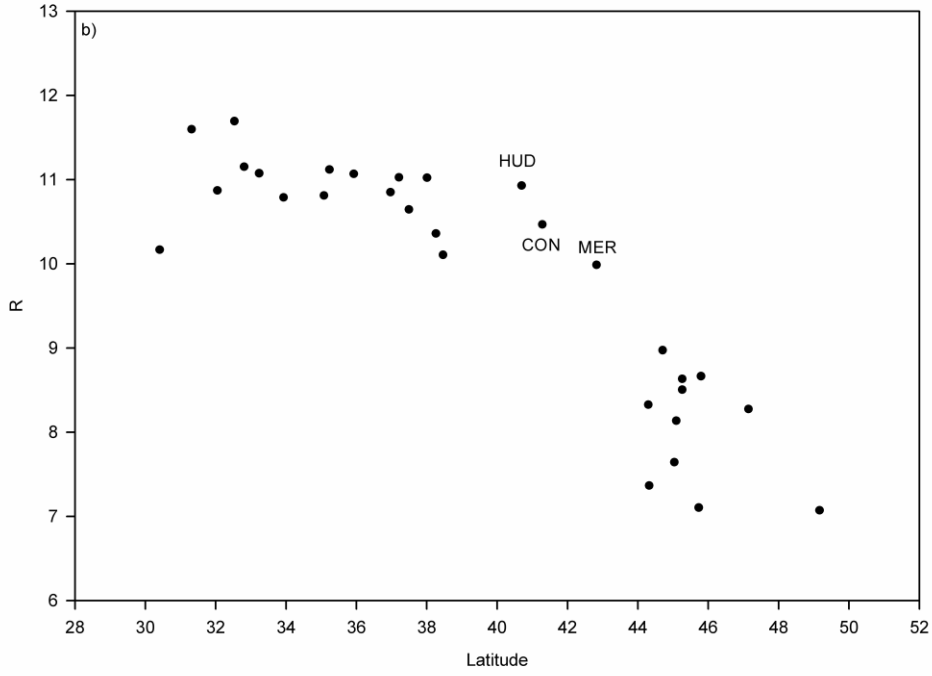
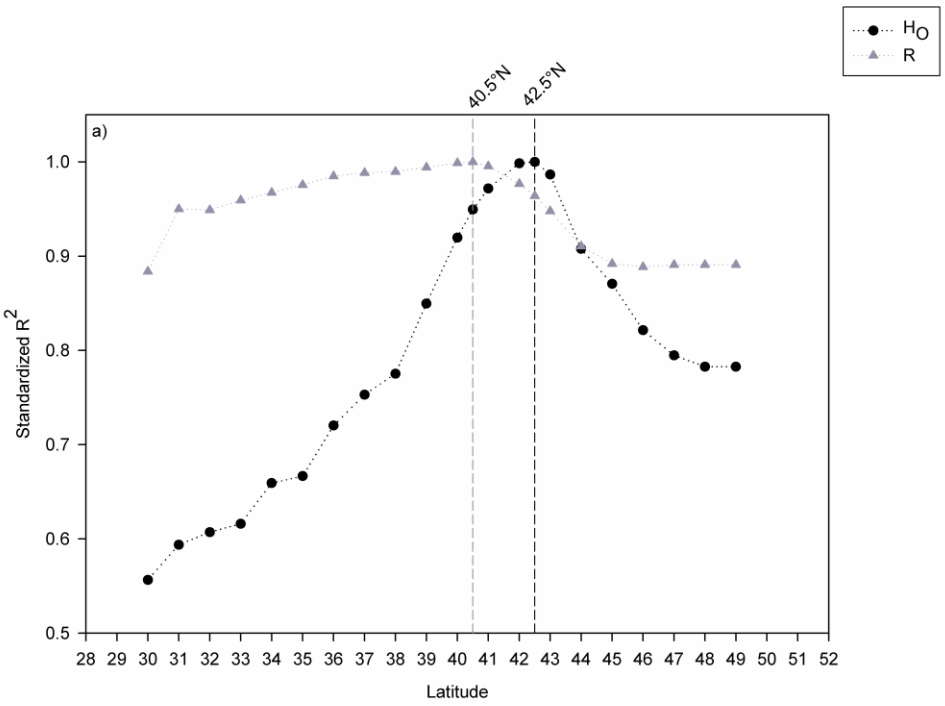
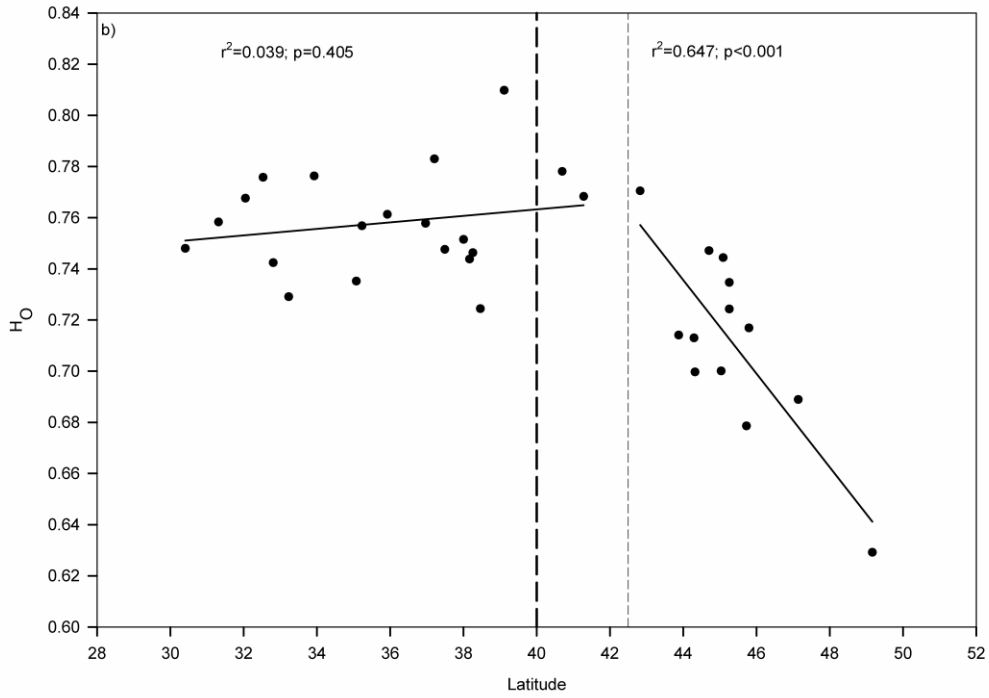
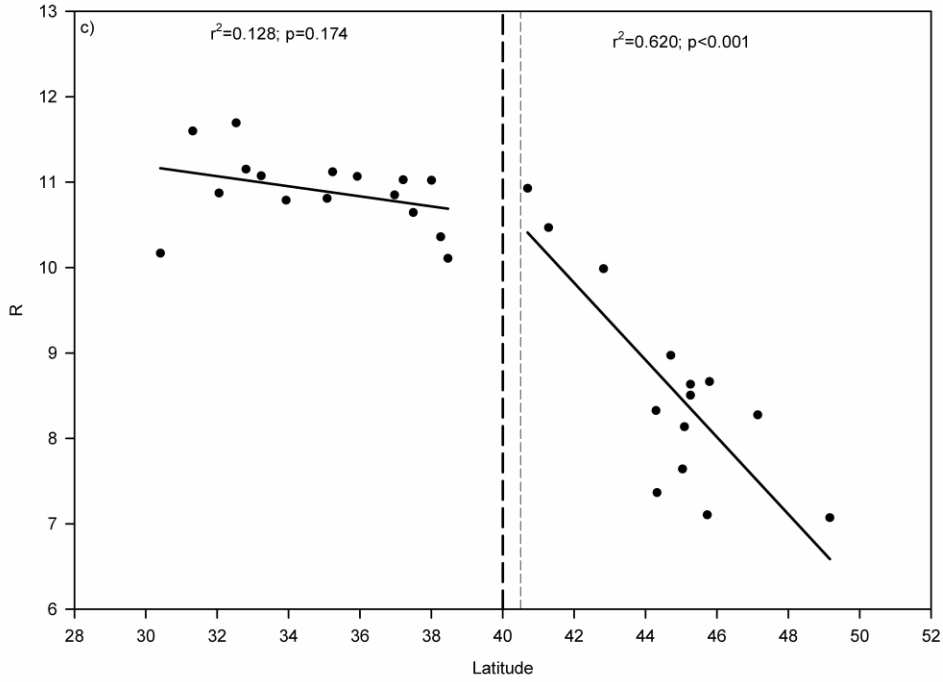

b)
HUD
CON
MER
R
Latitude
13
12
11
10
9
8
7
6
28
30
32
34
36
38
40
42
44
46
48
50
52

Figure 4.3 Geographic patterns of genetic diversity across the species range of American shad. a) Piecewise regression identified the latitudinal breakpoint for H_O (●) at 42.5°N, and for R (▲) at 40.5°N. Highly significant ($p<0.001$) decreases in genetic diversity were detected above the latitudinal breakpoints identified for H_O (b) and R (c). Long dashed black line indicates maximal extent of the Laurentide ice sheet during the Wisconsinan glacial event; short dashed grey line indicates latitudinal breakpoint for the measure of genetic diversity depicted.


H
O
R
40.5°N
42.5°N
a)
Standardized R2
1.0
0.9
0.8
0.7
0.6
0.5
28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52
Latitude


b)
r²=0.039; p=0.405
r²=0.647; p<0.001
H_O
Latitude


c)
r²=0.128; p=0.174
r²=0.620; p<0.001
R
Latitude

species range. Piecewise linear regression revealed that the greatest proportion of variation in H_O (r^2=0.67) was explained by a breakpoint at 42.5°N (Figure 4.3a), and grouped spawning runs from the Hudson and Connecticut rivers with those from rivers to the south, while the Merrimack River clustered with spawning runs to the north (Figure 4.3b). The greatest proportion of variation in R (r^2=0.86) was explained by a breakpoint at 40.5°N (Figure 4.3a), and grouped the Hudson, Connecticut, and Merrimack rivers with those from the formerly glaciated portion of the species range (Figure 4.3c). Linear regression revealed highly significant ($p<0.001$) declines of genetic diversity with increasing latitudes north of the latitudinal breakpoint for both H_O (r^2=0.65; Figure 3b) and R (r^2=0.62; Figure 4.3c).

4.3.3 Clinal Variation in Allele Frequencies

Linear regressions revealed evidence for significant ($p<0.05$) clinal variation for the frequency of the most common allele at 10 of 13 loci (Appendix 7). Asa4, Aa14, and Af20 exhibited non-significant clinal variation for the most common allele, but showed significant clinal variation for the second and third most common alleles (data not shown). A highly significant ($p<0.001$) linear cline was observed for the most common allele for five loci, and accounted for a substantial ($r^2 \geq 0.32$) proportion of variation in allele frequencies at these loci (Appendix 7). The most common allele for three of these loci (i.e. Asa16, AsaD042, and AsaD429) demonstrated increased frequency among formerly glaciated spawning runs, consistent with the influence of drift during range expansion (Appendix 7; panels d, l, and m, respectively).

4.3.4 GENETIC DIFFERENTIATION

An assessment of statistical power indicated that the microsatellite loci examined provided sufficient resolution to detect weak differentiation among spawning runs while maintaining the realized α-error near the intended level (0.05) for tests of genetic homogeneity. The probability of obtaining a significant ($p<0.05$) result in contingency tests among rivers with an F_{ST} of 0.001 was 1.00 (χ^2). Significant ($p<0.05$) genic differentiation between rivers was observed in 513 of 528 pairwise comparisons; non-significant comparisons occurred primarily among Chesapeake Bay drainages, and among semelparous spawning runs (Table 4.2).

Pairwise standardized estimates of genetic differentiation (F'_{ST}) among rivers ranged from -0.005-0.367 (F_{ST}= -0.001-0.108) (Appendix 8), and a multilocus standardized estimate of global differentiation was 0.271 (F_{ST}=0.044). Genetic differentiation among U.S. spawning runs was significantly weaker (mean F'_{ST}=0.044±0.031) than differentiation among Canadian spawning runs (mean F'_{ST}=0.157±0.067) (pairwise t-test; $p<0.001$). Non-significant ($p>0.05$) genetic differentiation (F_{ST}) was observed among most pairwise comparisons of neighbouring drainages in the U.S. (17/20), as well as among most rivers tributary to Chesapeake Bay (18/21), and among semelparous spawning runs (11/21) (Appendix 8). Hierarchical AMOVA revealed a highly significant ($p<0.001$) proportion of genetic variance partitioned among rivers, and among individuals within rivers (Table 4.3). A non-significant ($p>0.05$) component of variation was observed among temporal replicates, and suggested temporally stable population structure among spawning runs.

4.3.5 Relationships among Spawning Runs

Principal coordinates analysis revealed three factors which cumulatively explained 84.80% of the variation in genetic distance (D_A) among all spawning runs (Figure 4.4a). The first axis explained 58.32% of this variation, and linear regression revealed a highly significant (r^2=0.85; p<0.001) relationship between PCA axis-1 scores and latitude (Figure 4.4b). Canadian spawning runs exhibited greater variation along axis-2 than U.S. spawning runs (Figure 4.4a), but the factor(s) underlying this component of variation are uncertain.

4.3.6 Population Structure

While BAPS provided highly significant (p<0.001) support for eight clusters as the most likely number of genetically distinguishable groups among the spawning runs assessed, the maximum value of lnPr(X|*K*) using STRUCTURE was observed at *K*=10 (Figure 4.5a). However, the value of lnPr (X|*K*) for K=8 was not appreciably different than for K=10, and a decreased value (and greater variation among iterations) of lnPr(X|*K*) was observed for K=9 (Figure 4.5a); an indication that the true number of clusters in the data set has been surpassed (Pritchard et al. 2000). Cumulatively, these results suggested the presence of eight genetically distinguishable spawning groups among the 33 spawning runs examined in this study.

The eight clusters identified using BAPS and STRUCTURE were identical in composition and corresponded to the following regions: i) St. Lawrence River, ii) Gulf of St. Lawrence (Miramichi River and River Phillip), iii) Atlantic coast of Nova Scotia (Musquodoboit, LaHave, and Annis Rivers), iv) Annapolis and Gaspereau Rivers, v)

Table 4.2 Probability values for pairwise tests of genic heterogeneity among U.S. shad spawning runs. Instances of non-significant ($p>0.05$) genic heterogeneity are in bold.

	STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD
STL	.														
MR	0.000	.													
RP	0.000	0.000	.												
MUS	0.000	0.000	0.000	.											
LH	0.000	0.000	0.000	0.027	.										
AR	0.000	0.000	0.000	0.000	0.000	.									
ANN	0.000	0.000	0.000	0.000	0.000	0.000	.								
GR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.							
KR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.						
SHU	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.					
RH	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.				
SJR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.			

Table 4.2 *continued*

	STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD
MER	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.		
CON	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.	
HUD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.
DEL	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
NAN	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.018
SUS	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
PAT	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
POT	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RAPP	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
YOR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
JAM	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ROA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
TAR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.2 *continued*

	STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD
NEU	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
CF	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
WACC	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
COOP	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ED	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SAV	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ALT	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
STJ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.2 *continued*

	DEL	NAN	SUS	PAT	POT	RAPP	YOR	JAM	ROA	TAR	NEU	CF	WACC	COOP	ED
DEL	.														
NAN	0.025	.													
SUS	0.000	0.023	.												
PAT	0.002	**0.188**	0.000	.											
POT	0.001	**0.214**	0.000	0.010	.										
RAPP	0.002	0.048	0.000	**0.064**	**0.130**	.									
YOR	**0.056**	**0.279**	0.000	**0.106**	0.000	0.002	.								
JAM	0.000	**0.113**	0.000	0.000	0.008	0.019	0.000	.							
ROA	0.033	0.028	0.000	0.013	**0.169**	0.006	0.000	0.000	.						
TAR	0.003	**0.099**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.					
NEU	0.000	**0.067**	0.000	0.004	0.000	0.000	0.001	0.000	0.013	0.031	.				
CF	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.			
WACC	0.000	0.014	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.		

Table 4.2 *continued*

	DEL	NAN	SUS	PAT	POT	RAPP	YOR	JAM	ROA	TAR	NEU	CF	WACC	COOP	ED
COOP	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	**0.417**	.	
ED	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	**0.059**	0.003	.
SAV	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000	0.022	**0.258**	**0.149**
ALT	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.003	0.000
STJ	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

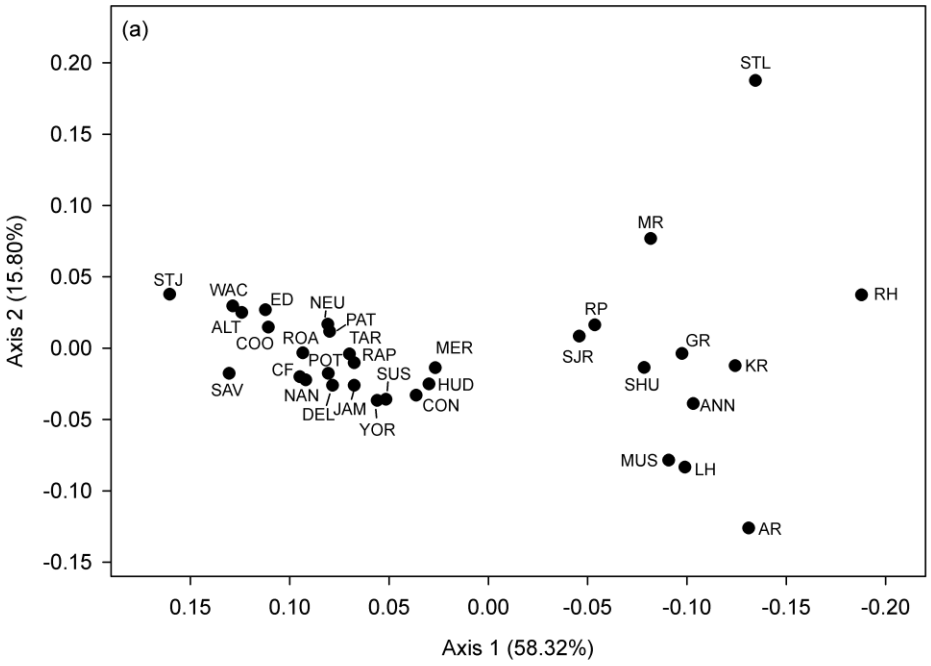
Table 4.2 *continued*

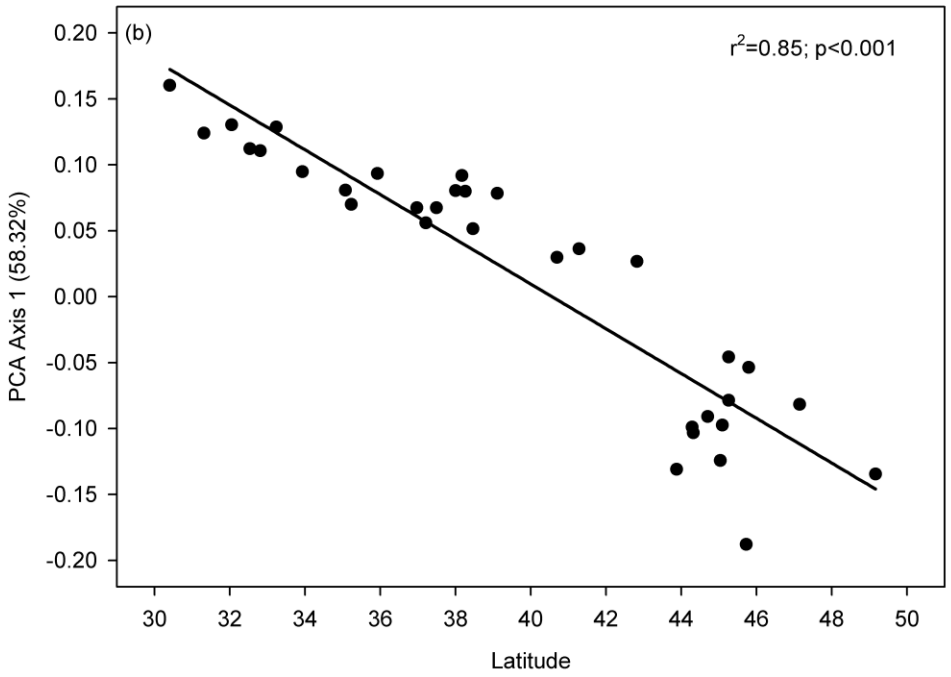
	SAV	ALTA	STJ
SAV	.		
ALTA	0.028	.	
STJ	0.000	0.000	

Table 4.3 Analysis of molecular variance among rivers, among temporal collections within river, and among clusters identified using Bayesian clustering analyses.

Variance component	d.f.	Sum of squares	% total variance	F-statistic	P-value
Among rivers	14	1158.09	4.98	0.050	0.000
Among collections within rivers	22	50.80	-0.27	-0.003	1.000
Among individuals within collections	5951	24692.05	95.29	0.047	0.000
Among clusters	8	1387.16	3.95	0.039	0.000
Among rivers within clusters	24	186.59	0.47	0.005	0.000
Among individuals within rivers	8707	36722.96	95.59	0.044	0.000

Figure 4.4 Genetic affinities among shad spawning runs using genetic distance (D_A; Nei et al. 1983): (a) the distribution of spawning runs along the first axis (axis order inverted to maintain consistency with latitude) in principal coordinates analysis (PCA) was broadly consistent with their geographic distribution; (b) linear regression revealed a strong relationship between PCA axis-1 scores and latitude of the spawning run assessed.

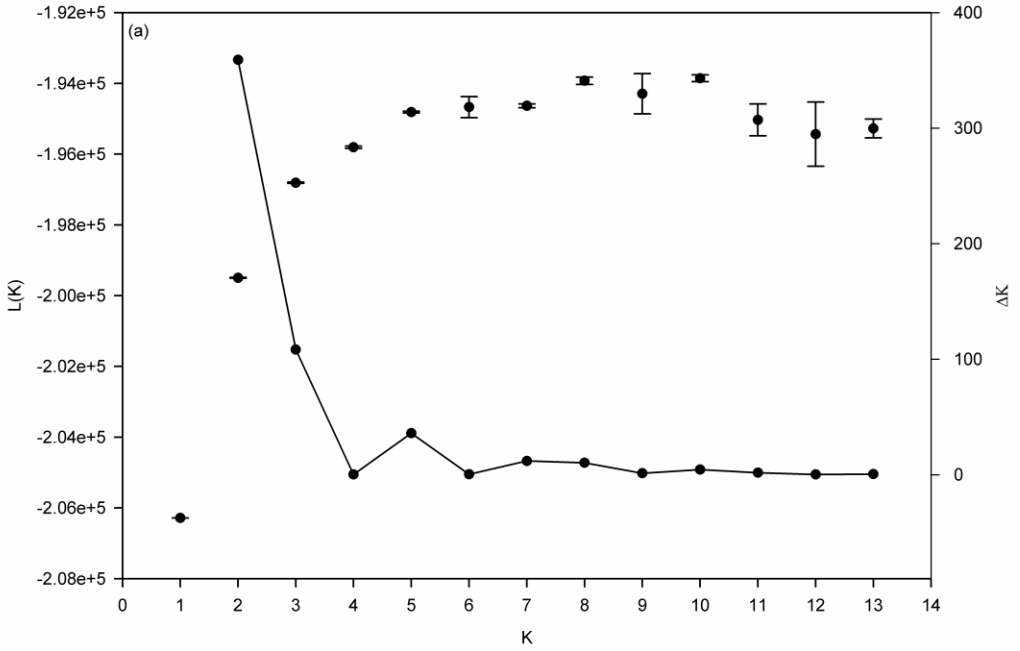

(a)
STL
MR
RH
STJ
WAC
ED
NEU
PAT
ALT
COO
ROA
TAR
RAP
MER
RP
SJR
GR
KR
POT
CF
SUS
HUD
SAV
NAN
DEL
JAM
CON
YOR
SHU
ANN
MUS
LH
AR
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
Axis 2 (15.80%)
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.20
Axis 1 (58.32%)

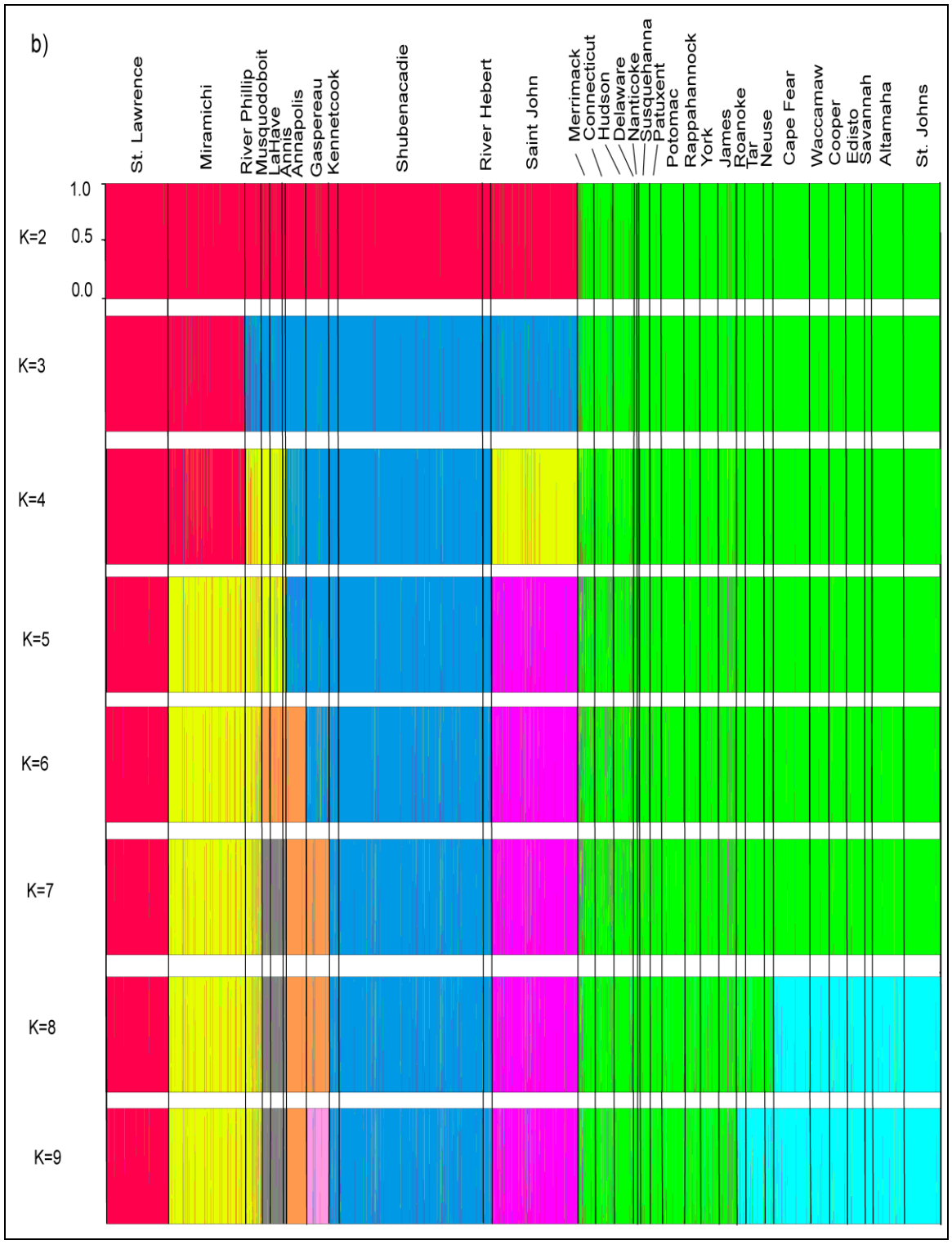

(b)
r2=0.85; p<0.001
PCA Axis 1 (58.32%)
0.20
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
-0.20
30
32
34
36
38
40
42
44
46
48
50
Latitude

Upper Bay of Fundy (River Hebert, Kennetcook River, and Shubenacadie River), vi) Saint John River, vii) iteroparous U.S. (Merrimack, Connecticut, Hudson, Delaware, Nanticoke, Susquehanna, Patuxent, Potomac, Rappahannock, York, James, Roanoke, Tar, Neuse Rivers), and viii) semelparous U.S. (Cape Fear, Waccamaw, Cooper, Edisto, Savannah, Altamaha, St. John's Rivers) (Figure 4.5b).

Further investigation of the eight clusters using hierarchical STRUCTURE analyses (Vaha et al. 2007) revealed that the two rivers in the Annapolis-Gaspereau cluster each constituted a genetically distinguishable group, and re-analysis of this cluster using BAPS corroborated this result. Hierarchical STRUCTURE analyses failed to detect further population structure within any of the remaining clusters, as the maximum value of $\ln Pr(X|K)$ within each cluster was observed at K=1 (data not shown); hierarchical BAPS analyses confirmed these observations. Cumulatively, these results suggest that there are nine genetically distinguishable groups of shad across the species' range. The specification of nine groups in BAPS admixture analysis altered the composition of two clusters, as the Roanoke, Tar and Neuse rivers shifted from the iteroparous U.S. cluster to the semelparous U.S. cluster (Figure 4.5b). Despite this alteration to cluster composition, the designation 'semelparous U.S. cluster' is retained throughout the remainder of this chapter to avoid confusion in terminology. A highly significant (AMOVA; $p<0.001$) component of variation was observed among clusters (3.95%), and was similar to that observed among rivers when not grouped into clusters (4.98%; Table 4.3). A highly significant ($p<0.001$) proportion of variation was also revealed among rivers within clusters (0.47%), but was considerably less than that

Figure 4.5 Assessment of number of American shad populations within the U.S. portion of the species range. (a) Bayesian inference of number of clusters (K) among 33 U.S. rivers using plateau of log probability of data L(K) (●; Pritchard et al. 2000), and ΔK (★; Evanno et al. 2005); (b) estimated population structure inferred from admixture analysis for K=9 clusters identified in Bayesian analyses. Individuals are represented by a thin vertical line which is partitioned into K-coloured segments representing an individual's estimated membership fractions from each of the identified clusters. Black lines separate individuals from different collections (labeled above).


(a)
-1.92e+5
-1.94e+5
-1.96e+5
-1.98e+5
-2.00e+5
-2.02e+5
-2.04e+5
-2.06e+5
-2.08e+5
L(K)
400
300
200
100
0
ΔK
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
K


b)
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehanna
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns
1.0
0.5
0.0
K=2
K=3
K=4
K=5
K=6
K=7
K=8
K=9

observed among clusters. This result was not surprising given the significant genic differentiation detected among most rivers within the nine clusters.

Estimates of ΔK revealed the largest increase in the likelihood of the number of clusters at K=2 (Figure 4.5a), and suggested that Canadian spawning runs comprised one cluster, while those in the U.S. comprised the other (Figure 4.5b). However, hierarchical analyses using STRUCTURE and BAPS conducted independently for each of these clusters identified the same nine clusters as described above (data not shown).

BAPS admixture analysis revealed relatively few individuals (n=133; 3.0%) with significantly ($p<0.05$) admixed ancestry. The Gulf of St. Lawrence cluster contained the greatest percentage of individuals with significantly admixed ancestry (9.51%), followed by the iteroparous U.S. cluster (9.33%), Gaspereau River (8.10%), Saint John River (7.62%), the upper Bay of Fundy cluster (6.41%), Atlantic Nova Scotia (5.38%), the semelparous U.S. cluster (4.08%), Saint Lawrence River (1.16%), and Annapolis River (0%). Among spawning runs, River Phillip contained the greatest percentage of individuals with significant ($p<0.05$) admixture (~22%), followed by the Nanticoke and Roanoke Rivers (~16% each). Remaining spawning runs exhibited a comparatively low percentage of individuals with significant admixture (mean=6.49±0.04%) from clusters other than the one to which they were assigned (Table 4.4).

4.3.7 BARRIER ANALYSIS

Using multilocus genotypes, BARRIER identified three discontinuities in gene flow across the species range that were supported by ≥7 loci. The first consensus barrier (supported by 10 loci) was placed between River Hebert and the Saint John River, and

Table 4.4 BAPS admixture proportions showing relative contributions of each of nine identified clusters to spawning run admixture. Values in bold are proportion of admixture from cluster of origin.

Pop	St. Lawrence	Gulf of St. Lawrence	Atlantic coast of Nova Scotia	Annapolis	Gaspereau	Upper Bay of Fundy	Saint John	Iteroparous U.S.	Semelparous U.S.
STL	**0.9884**	0.0012	0.0002	0.0020	0.0042	0.0023	0.0007	0.0006	0.0005
MR	0.0392	**0.9320**	0.0054	0.0034	0.0094	0.0040	0.0022	0.0032	0.0014
RP	0.0138	**0.7776**	0.0865	0.0216	0.0044	0.0387	0.0530	0.0008	0.0035
MUS	0.0000	0.0049	**0.9800**	0.0016	0.0000	0.0000	0.0135	0.0000	0.0000
LH	0.0081	0.0146	**0.9301**	0.0066	0.0025	0.0112	0.0112	0.0131	0.0026
AR	0.0147	0.0229	**0.9253**	0.0000	0.0000	0.0371	0.0000	0.0000	0.0000

Table 4.4 *continued*

Pop	St. Lawrence	Gulf of St. Lawrence	Atlantic coast of Nova Scotia	Annapolis	Gaspereau	Upper Bay of Fundy	Saint John	Iteroparous U.S.	Semelparous U.S.
ANN	0.0000	0.0000	0.0000	**1.0000**	0.0000	0.0000	0.0000	0.0000	0.0000
GR	0.0040	0.0080	0.0052	0.0243	**0.9190**	0.0323	0.0045	0.0008	0.0019
KR	0.0082	0.0180	0.0122	0.0236	0.0194	**0.9010**	0.0126	0.0004	0.0046
SHU	0.0127	0.0097	0.0052	0.0115	0.0080	**0.9417**	0.0048	0.0034	0.0029
RH	0.0127	0.0129	0.0047	0.0178	0.0707	**0.8780**	0.0000	0.0002	0.0031
SJR	0.0131	0.0193	0.0055	0.0127	0.0068	0.0058	**0.9238**	0.0053	0.0076
MER	0.0161	0.0110	0.0127	0.0195	0.0049	0.0260	0.0197	**0.8765**	0.0135

Table 4.4 *continued*

Pop	St. Lawrence	Gulf of St. Lawrence	Atlantic coast of Nova Scotia	Annapolis	Gaspereau	Upper Bay of Fundy	Saint John	Iteroparous U.S.	Semelparous U.S.
CON	0.0074	0.0090	0.0070	0.0091	0.0411	0.0124	0.0305	**0.8606**	0.0229
HUD	0.0056	0.0120	0.0015	0.0010	0.0036	0.0040	0.0147	**0.9488**	0.0087
DEL	0.0008	0.0208	0.0048	0.0004	0.0296	0.0020	0.0000	**0.8800**	0.0616
NAN	0.0017	0.0000	0.0000	0.0000	0.0675	0.0358	0.0025	**0.8333**	0.0592
SUS	0.0059	0.0000	0.0002	0.0085	0.0063	0.0004	0.0000	**0.9630**	0.0157
PAT	0.0100	0.0000	0.0000	0.0070	0.0079	0.0020	0.0116	**0.9393**	0.0221
POT	0.0082	0.0045	0.0031	0.0046	0.0039	0.0001	0.0022	**0.9372**	0.0362

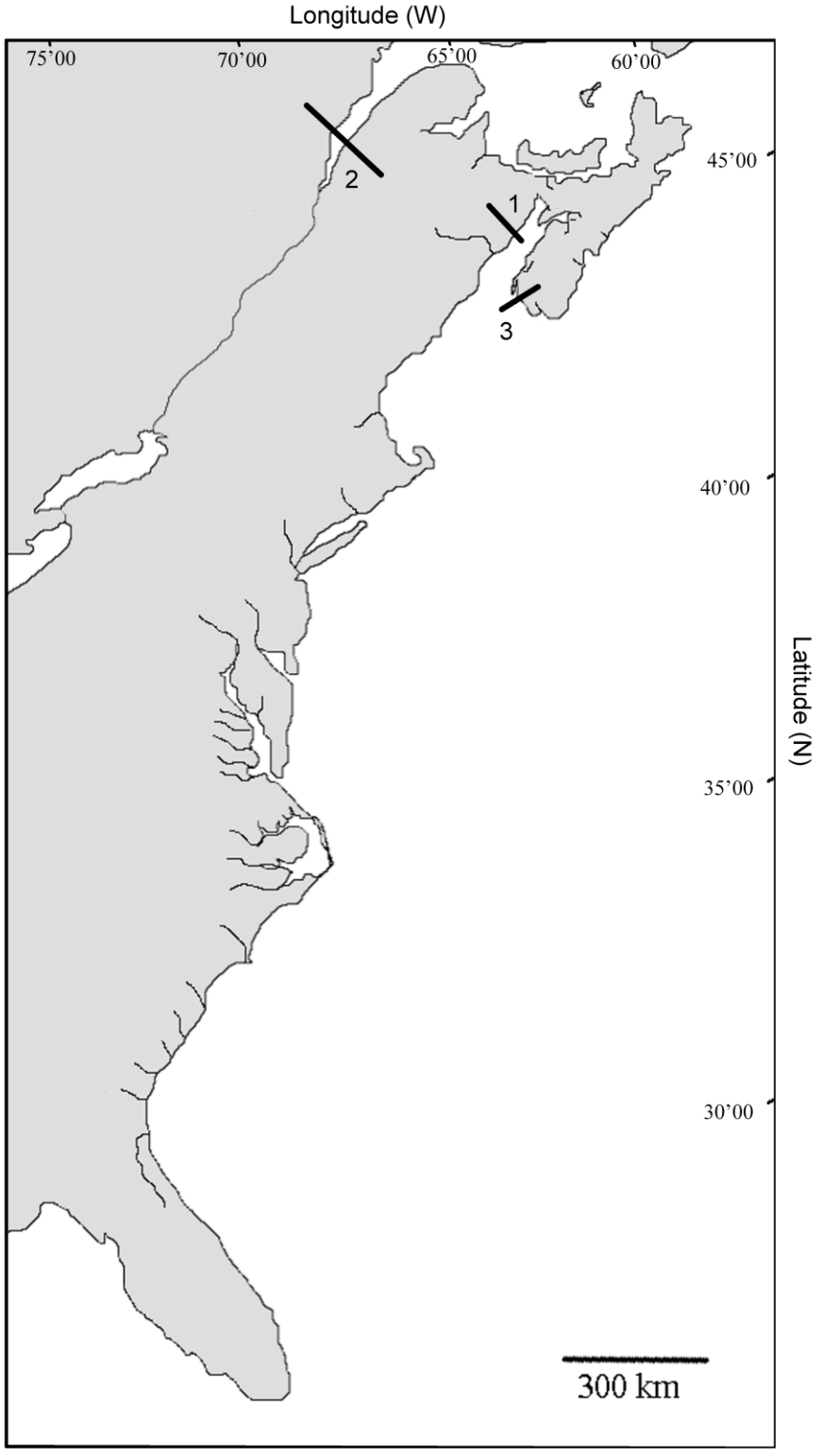
Table 4.4 *continued*

Pop	St. Lawrence	Gulf of St. Lawrence	Atlantic coast of Nova Scotia	Annapolis	Gaspereau	Upper Bay of Fundy	Saint John	Iteroparous U.S.	Semelparous U.S.
RAPP	0.0082	0.0055	0.0045	0.0149	0.0064	0.0072	0.0025	**0.9069**	0.0439
YOR	0.0063	0.0006	0.0062	0.0060	0.0055	0.0123	0.0103	**0.8888**	0.0641
JAM	0.0237	0.0025	0.0029	0.0025	0.0122	0.0037	0.0479	**0.8822**	0.0224
ROA	0.0053	0.0084	0.0127	0.0120	0.0192	0.0059	0.0124	0.0847	**0.8394**
TAR	0.0053	0.0088	0.0065	0.0116	0.0063	0.0089	0.0024	0.0321	**0.9180**
NEU	0.0045	0.0004	0.0121	0.0068	0.0123	0.0021	0.0185	0.0294	**0.9138**
CF	0.0034	0.0059	0.0031	0.0086	0.0048	0.0047	0.0116	0.0057	**0.9521**

Table 4.4 *continued*

Pop	St. Lawrence	Gulf of St. Lawrence	Atlantic coast of Nova Scotia	Annapolis	Gaspereau	Upper Bay of Fundy	Saint John	Iteroparous U.S.	Semelparous U.S.
WACC	0.0071	0.0026	0.0000	0.0020	0.0012	0.0005	0.0049	0.0058	**0.9759**
COOP	0.0052	0.0073	0.0043	0.0004	0.0050	0.0001	0.0000	0.0010	**0.9766**
ED	0.0052	0.0082	0.0046	0.0060	0.0028	0.0003	0.0001	0.0000	**0.9727**
SAV	0.0121	0.0000	0.0021	0.0000	0.0000	0.0000	0.0023	0.0036	**0.9800**
ALT	0.0033	0.0122	0.0000	0.0042	0.0056	0.0015	0.0051	0.0028	**0.9652**
STJ	0.0001	0.0005	0.0008	0.0000	0.0000	0.0013	0.0013	0.0000	**0.9960**

Figure 4.6 Major discontinuities in gene flow suggested by BARRIER using consensus barriers (>7 loci), with rank order given in Arabic numerals. The first barrier was supported by 10 loci, the second barrier by 7 loc, and the third by 8 loci. Six additional barriers were identified, but were weakly supported (< 7 loci), and were restricted to the Canadian portion of the species range.


Longitude (W)
75’00
70’00
65’00
60’00
45’00
40’00
35’00
30’00
Latitude (N)
1
2
3
300 km

grouped the Saint John River spawning run with those in the U.S (Figure 4.6). The second consensus barrier (7 loci) isolated the St. Lawrence River from remaining spawning runs (Figure 4.6). The third consensus barrier (8 loci) separated the Annis and Annapolis rivers, and together with the first barrier, effectively isolated Bay of Fundy drainages from the remainder of the species' range (Figure 4.6). Although additional barriers were identified, they were only weakly supported (<7 loci). The fourth barrier (6 loci) separated the Annapolis and Gaspereau rivers, and the fifth barrier (3 loci) separated the Shubenacadie River from River Hebert. These five barriers are identical to, and of the same rank order as, those identified through the examination of the Canadian portion of the species range (Chapter 3). Although 10 barriers were identified in total, barriers 1-9 were positioned among Canadian rivers, while the 10th barrier was positioned between the Hudson River and the Delaware River, consistent with the LGM. No additional discontinuities in gene flow were observed among U.S. rivers.

4.3.8 Isolation by Distance

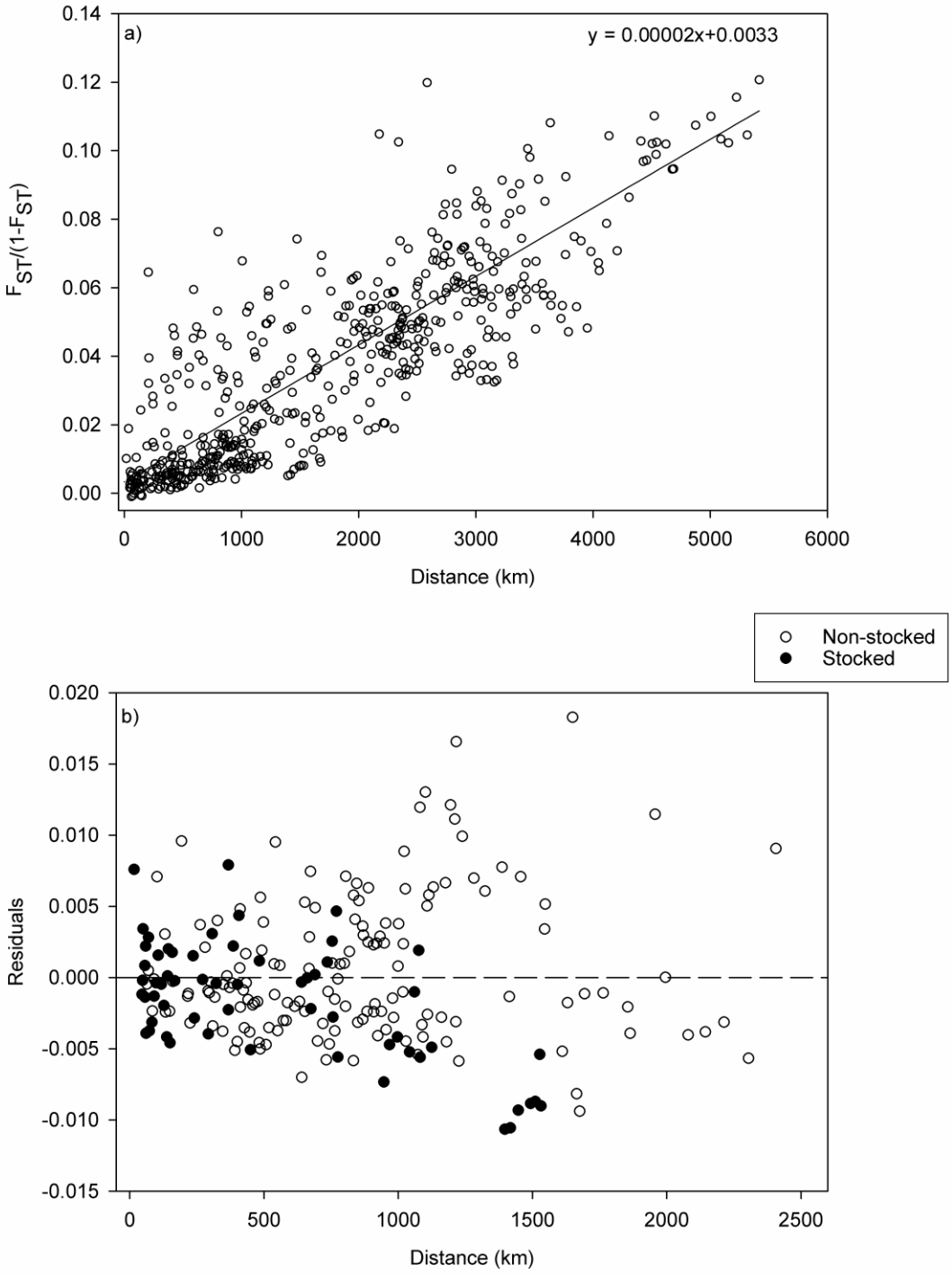
Mantel tests revealed a highly significant ($p<0.001$; $r=0.84$) pattern of IBD among all spawning runs from across the species range (Figure 4.7a). Highly significant IBD ($p<0.001$; $r=0.71$) was also observed among U.S. spawning runs. Examination of historical shad stocking records revealed evidence of stock transfers for 59 pairs of spawning populations (i.e. donor-recipient spawning stocks) from the iteroparous portion of the species' range. Examination of the distribution of residuals from a linear regression of the IBD pattern observed among U.S. spawning runs revealed 39 pairs of spawning populations that exhibited negative residuals (i.e. lower genetic differentiation than what might be expected based on geographic distance alone) (Figure 4.7b). Among

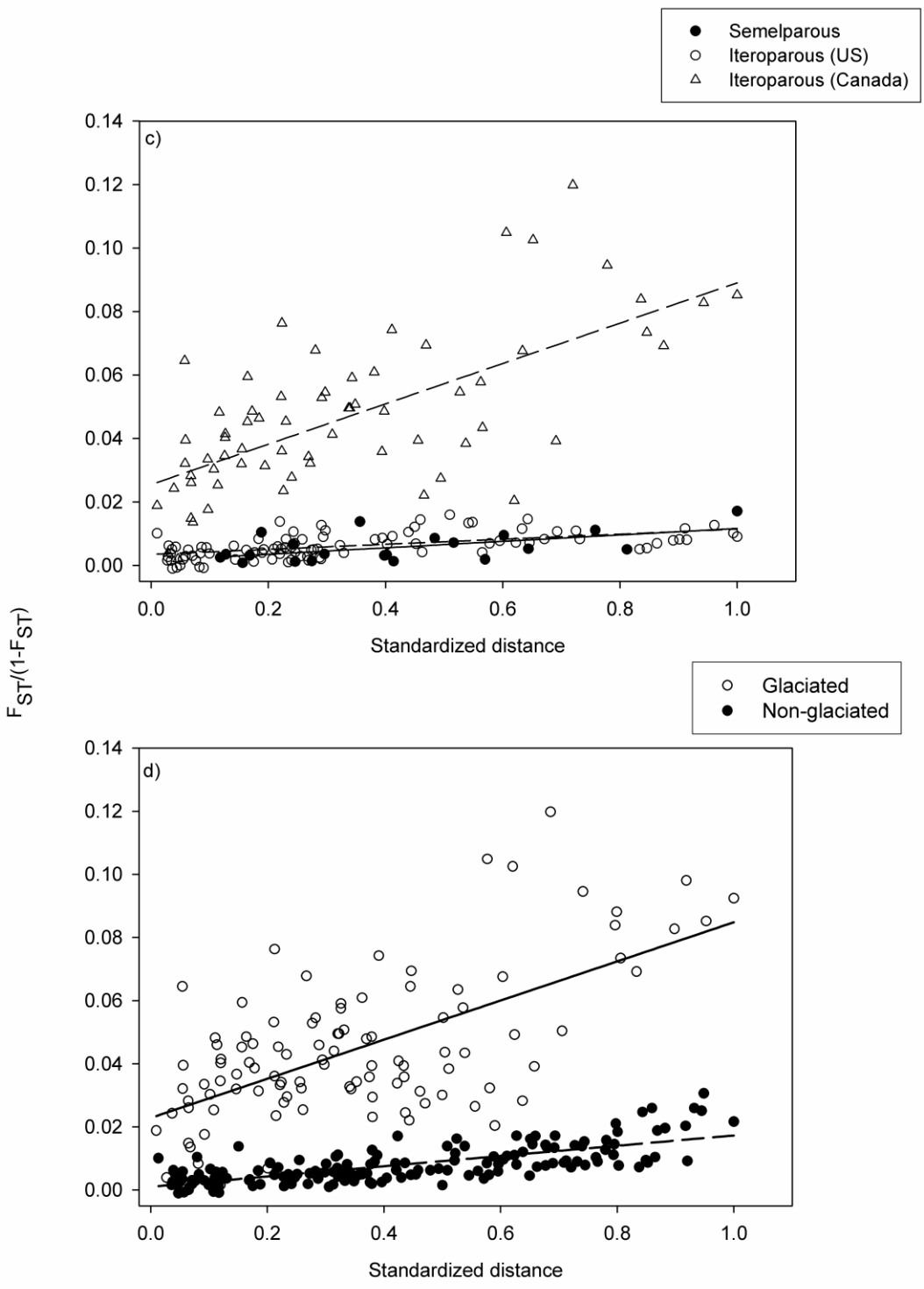
the 59 pairs of spawning runs, a general decline in the value of residuals was observed for spawning run pairs spaced greater than ~500 km apart.

Separate analyses of iteroparous and semelparous populations of shad revealed a highly significant pattern of IBD among iteroparous spawning runs ($p<0.001$; $r=0.77$), but only a marginally significant IBD pattern among semelparous spawning runs ($p=0.044$; $r=0.55$). ANCOVA revealed heterogeneity of slopes between iteroparous and semelparous spawning runs (i.e. significant life history by covariate interaction; $p<0.005$), and suggested different spatial patterns of population structure in the two portions of the species range characterized by alternate reproductive strategies (Figure 4.7c). However, these patterns were largely driven by the inclusion of Canadian (iteroparous) spawning runs, as ANCOVA revealed homogeneity of slopes (non-significant life history by covariate interaction; $p=0.44$) between U.S. iteroparous and semelparous spawning runs, but heterogeneity of slopes for comparisons of Canadian spawning runs with both U.S iteroparous and semelparous spawning runs (significant life history by covariate interaction; $p<0.001$ and $p<0.02$, respectively) (Figure 4.7c). The similar IBD patterns observed between U.S. iteroparous and semelparous spawning runs suggests that life history variation is not responsible for different patterns of population structure observed across the species' range.

Alternatively, different management strategies in Canada (no stocking) and the U.S. (iteroparous) (stocking) may have lead to the different patterns of population structure, as stock transfers appear to have decreased the level of genetic differentiation below that expected based on geographic distance alone. However, the different patterns of population structure observed across the species range may also reflect alternate

Figure 4.7 Relationship between pairwise genetic differentiation (F_{ST}) and geographic distance; (a) pairwise comparisons among all rivers; (b) regression residuals for IBD among U.S. spawning runs showing stocked (●) and non-stocked (○) spawning population pairs; (c) linear regression of genetic differentiation on standardized geographic distance among Canadian (△) and U.S. (○) iteroparous, and semelparous (●) spawning runs; (d) linear regression of genetic differentiation on standardized geographic distance among formerly glaciated (○) and non-glaciated (●) rivers. ANCOVA revealed different patterns of isolation by distance between Canadian and U.S. (iteroparous and semelparous) spawning runs, but homogeneity of IBD slopes between U.S. iteroparous and semelparous spawning runs. Heterogeneity of IBD slopes was observed between spawning runs from formerly glaciated and non-glaciated rivers.


a)
y = 0.00002x+0.0033
$F_{ST}/(1-F_{ST})$
Distance (km)
Non-stocked
Stocked
b)
Residuals
Distance (km)


Semelparous
Iteroparous (US)
Iteroparous (Canada)
c)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0.0
0.2
0.4
0.6
0.8
1.0
Standardized distance
$F_{ST}/(1-F_{ST})$
Glaciated
Non-glaciated
d)
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0.0
0.2
0.4
0.6
0.8
1.0
Standardized distance

glacial histories. Categorization of rivers into formerly glaciated and non-glaciated groups largely overlapped the classification of these rivers into Canadian and U.S. spawning runs. ANCOVA revealed a highly significant ($p<0.001$) glacial history by covariate interaction, and heterogeneity of slopes between formerly glaciated and non-glaciated rivers (Figure 4.7d).

4.3.9 Contributions to Genetic Diversity

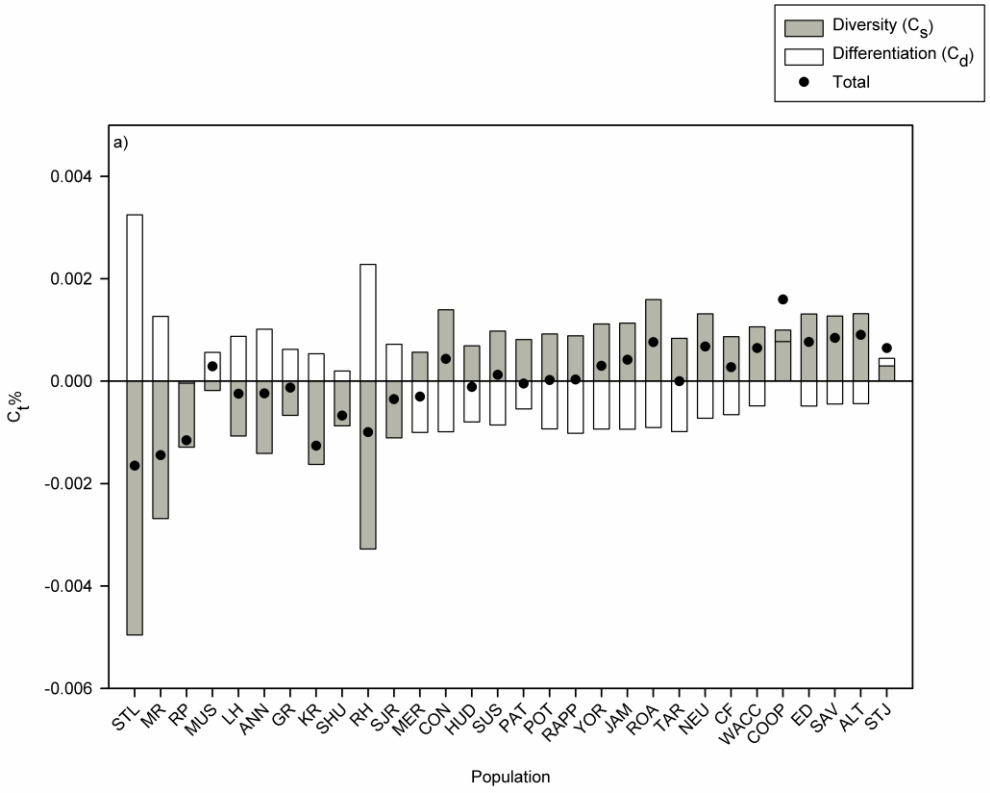
Analyses of the relative contributions of each spawning run to species level heterozygosity suggested that most Canadian rivers (10 of 11) contributed below average, and most U.S. rivers (15 of 19) above average levels of diversity to global H_O (Figure 4.8a). Pairwise t-tests (SYSTAT 11; SPSS 2000) confirmed that U.S. spawning runs contributed a significantly ($p<0.001$) greater proportion of heterozygosity to global H_O, and exhibited significantly ($p<0.001$) greater variation in their relative contributions to global H_O, than Canadian spawning runs. The partitioning of this relative contribution into components reflecting within population diversity (C_s) and among population differentiation (C_d) helped to explain this pattern, and revealed reciprocal trends between Canadian and U.S. rivers (Figure 4.8a). Specifically, below average contributions by Canadian rivers to global H_O were due to significantly ($p<0.001$) lower levels of within population diversity detected among these stocks (relative to their U.S. counterparts; Figure 4.8a). Canadian rivers also exhibited significantly greater ($p<0.001$) levels of among population differentiation than U.S. spawning runs (Figure 4.8a).

Among all rivers, the Cooper spawning run contributed the greatest to global H_O, and exhibited a greater level of differentiation than neighbouring rivers (Table 4.5, Figure 4.8a). At the northern edge of the species range, the St. Lawrence spawning run

contributed the least to global $H_{O,}$ and possessed the lowest level of within population diversity, but highest level of among population differentiation of any stock examined. A similar pattern was revealed at the southern edge of the species range, where the Saint Johns River exhibited reduced levels of within population diversity, but increased levels of differentiation compared to neighbouring rivers. Spawning runs near the centre of the species distribution (i.e. Patuxent, Potomac, Rappahannock rivers) contributed average levels of heterozygosity to global H_O, and exhibited levels of within population diversity and among population differentiation on par with neighbouring spawning runs (Figure 4.8a).

Analyses of the relative contribution of each spawning run to species allelic richness revealed that individual rivers contributed comparatively little diversity to global R (Figure 4.8b). This may be due to the proportion of alleles shared among populations (Appendix 6), and the low number and rare frequency (<0.05) of private alleles detected. Nonetheless, U.S. spawning runs contributed a significantly ($p<0.001$; pairwise t-test) greater proportion of alleles to global R, and exhibited significantly ($p<0.001$) greater variation in their relative contributions to global R, than Canadian rivers. Partitioning of the relative contribution of each spawning run into a component reflecting allelic diversity (C_{rs}; the number of alleles observed within populations) and allelic differentiation (C_{rd}; alleles not present in other populations) revealed patterns broadly concordant with that observed for heterozygosity (Figure 4.8b). A reciprocal pattern similar to that observed for within population heterozygosity was observed among Canadian and U.S. spawning runs for within population levels of allelic diversity, as Canadian rivers exhibited significantly ($p<0.001$) lower levels of within population allelic

Figure 4.8 Relative contribution (●) of each population to species level a) total H_O (C_t%), and b) total R (C_{rt}%). Values above the horizontal line indicate populations that contribute more than average to genetic diversity, while those below contribute less than average to genetic diversity. Relative contributions of each population were subdivided into a diversity and differentiation component (see text).


Diversity (C_s)
Differentiation (C_d)
Total
a)
0.004
0.002
0.000
-0.002
-0.004
-0.006
C_t%
STL MR RP MUS LH ANN GR KR SHU RH SJR MER CON HUD SUS PAT POT RAPP YOR JAM ROA TAR NEU CF WACC COOP ED SAV ALT STJ
Population

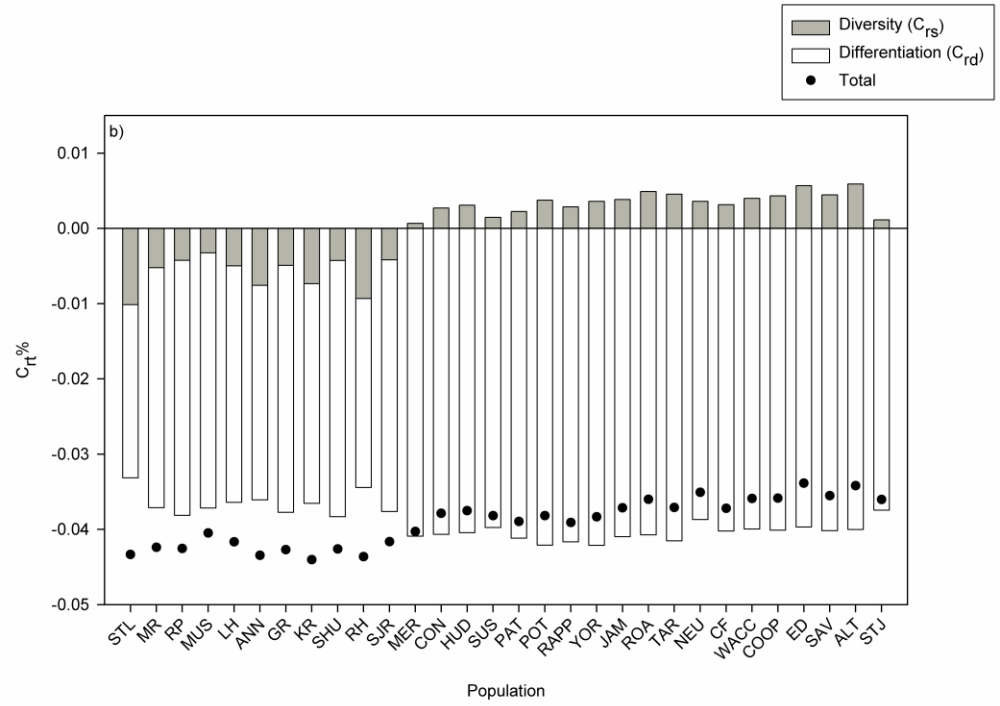

Diversity (C_{rs})
Differentiation (C_{rd})
Total
b)
0.01
0.00
-0.01
-0.02
-0.03
-0.04
-0.05
C_{rt}%
STL MR RP MUS LH ANN GR KR SHU RH SJR MER CON HUD SUS PAT POT RAPP YOR JAM ROA TAR NEU CF WACC COOP ED SAV ALT STJ
Population

Table 4.5 Multilocus estimates of genetic diversity for 33 spawning populations of American shad collected from across the species range. Unbiased heterozygosity (H_E; Nei 1978), observed heterozygosity (H_O), F_{IS} (Weir and Cockerham 1984), and allelic richness (R; standardized to N=32) are shown. Components of genetic diversity are as follows: C_t=the contribution of the *k*th population to total H_O; C_s=the contribution of the *k*th population to total H_O based on *k*'s own diversity; C_d= the contribution of the *k*th population to total H_O based on *k*'s differentiation from other populations; C_{rt}=the contribution of the *k*th population to total R; C_{rs}= the contribution of the *k*th population to total R based on *k*'s own R; C_{rd}= the contribution of the *k*th population to total R due to *k*'s allelic divergence or uniqueness.

Population	H_E	H_O	F_{IS}	R	C_t	C_s	C_d	C_{rt}	C_{rs}	C_{rd}
STL	0.6495	0.6292	0.031	7.07	-0.0016	-0.0050	0.0032	-0.0433	-0.0102	-0.0332
MR	0.7016	0.6889	0.018	8.27	-0.0014	-0.0027	0.0013	-0.0424	-0.0052	-0.0371
RP	0.7342	0.7169	0.024	8.67	-0.0012	-0.0013	0.0000	-0.0425	-0.0043	-0.0382
MUS	0.7578	0.7471	0.014	8.97	0.0003	-0.0002	0.0006	-0.0405	-0.0032	-0.0372
LH	0.7394	0.7130	0.036	8.33	-0.0002	-0.0011	0.0009	-0.0417	-0.0050	-0.0364

Table 4.6 *continued*

Population	H_E	H_O	F_{IS}	R	C_t	C_s	C_d	C_{rt}	C_{rs}	C_{rd}
AR	0.7309	0.7141	0.024	NA	NA	NA	NA	NA	NA	NA
ANN	0.7275	0.6997	0.038	7.37	-0.0002	-0.0014	0.0010	-0.0435	-0.0076	-0.0361
GR	0.7444	0.7444	0.000	8.14	-0.0001	-0.0007	0.0006	-0.0427	-0.0049	-0.0378
KR	0.7229	0.7001	0.032	7.64	-0.0013	-0.0016	0.0005	-0.0440	-0.0074	-0.0366
SHU	0.7432	0.7347	0.011	8.51	-0.0007	-0.0009	0.0002	-0.0426	-0.0043	-0.0383
RH	0.6859	0.6786	0.011	7.10	-0.0010	-0.0033	0.0023	-0.0436	-0.0093	-0.0345
SJR	0.7379	0.7243	0.019	8.63	-0.0004	-0.0011	0.0007	-0.0416	-0.0042	-0.0376
MER	0.7779	0.7705	0.010	9.99	-0.0003	0.0006	-0.0010	-0.0403	0.0006	-0.0409
CON	0.7944	0.7683	0.033	10.47	0.0004	0.0014	-0.0010	-0.0379	0.0027	-0.0407

Table 4.6 *continued*

Population	H_E	H_O	F_{IS}	R	C_t	C_s	C_d	C_{rt}	C_{rs}	C_{rd}
HUD	0.7773	0.7781	-0.001	10.93	-0.0001	0.0007	-0.0008	-0.0375	0.0031	-0.0405
DEL	0.8043	0.8098	-0.007	NA	NA	NA	NA	NA	NA	NA
NAN	0.7868	0.7438	0.057	NA	NA	NA	NA	NA	NA	NA
SUS	0.7848	0.7244	0.078	10.11	0.0001	0.0010	-0.0009	-0.0382	0.0015	-0.0398
PAT	0.7775	0.7463	0.040	10.36	0.0000	0.0008	-0.0005	-0.0389	0.0022	-0.0412
POT	0.7834	0.7515	0.041	11.02	0.0000	0.0009	-0.0009	-0.0382	0.0037	-0.0421
RAPP	0.7865	0.7476	0.050	10.65	0.0000	0.0009	-0.0010	-0.0391	0.0029	-0.0417
YOR	0.7865	0.7830	0.008	11.03	0.0003	0.0011	-0.0009	-0.0383	0.0036	-0.0421
JAM	0.7928	0.7578	0.044	10.85	0.0004	0.0011	-0.0009	-0.0371	0.0038	-0.0410

Table 4.6 *continued*

Population	H_E	H_O	F_{IS}	R	C_t	C_s	C_d	C_{rt}	C_{rs}	C_{rd}
ROA	0.7981	0.7613	0.047	11.07	0.0008	0.0016	-0.0009	-0.0360	0.0049	-0.0408
TAR	0.7820	0.7568	0.032	11.12	0.0000	0.0008	-0.0010	-0.0371	0.0045	-0.0415
NEU	0.7907	0.7352	0.071	10.81	0.0007	0.0013	-0.0007	-0.0351	0.0036	-0.0387
CF	0.7830	0.7763	0.009	10.79	0.0003	0.0009	-0.0007	-0.0372	0.0031	-0.0402
WACC	0.7860	0.7291	0.073	11.08	0.0006	0.0011	-0.0005	-0.0359	0.0040	-0.0400
COOP	0.7839	0.7424	0.053	11.15	0.0016	0.0010	0.0008	-0.0359	0.0043	-0.0401
ED	0.7908	0.7758	0.019	11.69	0.0008	0.0013	-0.0005	-0.0339	0.0057	-0.0397
SAV	0.7914	0.7676	0.030	10.87	0.0008	0.0013	-0.0004	-0.0355	0.0044	-0.0402
ALT	0.7940	0.7583	0.045	11.60	0.0009	0.0013	-0.0004	-0.0342	0.0059	-0.0400

Table 4.6 *continued*

Population	H_E	H_O	F_{IS}	R	C_t	C_s	C_d	C_{rt}	C_{rs}	C_{rd}
STJ	0.7679	0.7480	0.026	10.17	0.0006	0.0003	0.0004	-0.0360	0.0011	-0.0375

diversity than U.S. spawning runs. Canadian rivers also exhibited a significantly ($p<0.001$) greater level of among population allelic differentiation than U.S. rivers (Table 4.5; Figure 4.8b), suggesting that a higher proportion of alleles are shared among U.S. spawning runs than among Canadian spawning runs. As observed for H_O, allelic diversity was lowest and allelic differentiation highest at the northern edge of the species range. A similar pattern of reduced diversity and elevated differentiation was observed at the southern edge of the species range.

4.4 DISCUSSION

While a stepping stone model of population structure provides the foundation for the spatial distribution of neutral genetic variation in shad, the relative influences of historical demography, microevolutionary processes, and anthropogenic factors on contemporary patterns varies depending on the portion of the range considered. The spatial distribution of genetic variation among formerly glaciated spawning populations probably reflects the combined influences of stepwise post-glacial colonization and the effects of natural gene flow and genetic drift since that time. Although weak population structure observed among U.S. iteroparous spawning populations may reflect the combined effects of historical demography and the influence of stock transfers among spawning populations, weak genetic differentiation detected among semelparous spawning runs may be a consequence of elevated levels of gene flow in this portion of the species' range. The spatial patterns of genetic variation and population structure observed within each of these regions are discussed below.

4.4.1 SPATIAL TRENDS IN GENETIC DIVERSITY

This survey of the distribution of neutral genetic variation in shad revealed different spatial patterns of genetic diversity among spawning runs from non-glaciated and formerly glaciated portions of the species' range. Although no significant relationship was detected between genetic diversity and latitude south of the LGM (i.e. $\leq$40°N), the spatial distribution of allelic richness suggests a slow decline with latitude in this region. Conversely, highly significant ($p<0.001$) reductions in diversity with latitude were observed among spawning runs from formerly glaciated regions (Figure 4.3b; Figure 4.3c). These results are consistent with those for Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) (King et al. 2001), an anadromous species with similar distribution (Waldman et al. 2002), and suggests that sequential reductions in shad genetic diversity among formerly glaciated rivers are a consequence of a stepwise process of post-glacial range expansion, and are associated with successive population founder events.

Prior surveys of genetic variation in shad (reviewed in Nolan et al. 2003) failed to reveal reductions in diversity among spawning runs from formerly glaciated regions, but were restricted in scope to few markers, few sampled rivers, or a combination thereof. Further, most of those studies examined geographic variation in mtDNA, a single non-recombining molecular marker that does not provide independently replicated data for inference of population histories (Ballard and Whitlock 2004). Multiple unlinked microsatellites provide independent data about population histories, and can retain high levels of polymorphism in founder populations (Bonhomme et al. 2008). Although a prior assessment of shad using microsatellites detected little geographic variation in

genetic diversity (Waters et al. 2000), that study was limited in scope to few loci and few spawning runs near the center of the species range. This study examines data from 13 polymorphic loci and 33 spawning from across the species range, and highlights the importance of adequately sampling across species distributions before making inferences about range wide patterns of genetic variation.

Although genetic diversity was generally lower among spawning runs from formerly glaciated regions, the Hudson, Connecticut and Merrimack Rivers exhibited levels of observed heterozygosity and allelic richness similar to spawning runs from non-glaciated regions, and generated uncertainty about where the boundary between the different latitudinal trends observed across the species' range should be placed. The latitude at which the decline in genetic diversity began (H_O) or increased in rate (R) approximated the maximal extent of the Laurentide ice sheet during Wisconsinan glaciation (40°N). However, the precise location varied depending on the diversity measure considered. Declines in heterozygosity commenced at 42.5°N and grouped the Hudson and Connecticut rivers with those from non-glaciated regions (Figure 4.3b), while an increase in the rate of declines for allelic richness was more consistent with the LGM (i.e. 40.5°N; Figure 4.3c).

The discrepancy in the latitudinal breakpoint for these genetic diversity measures is interesting in that it suggests the possibility of a different rate of recovery for heterozygosity and allelic diversity from demographic events, and may reflect properties of the genetic diversity measures themselves. Reductions in population size experienced during a founder event or other population bottleneck are anticipated to exhibit a greater influence on allelic diversity than heterozygosity (Allendorf 1986). Specifically, allelic

diversity declines more rapidly than heterozygosity during a bottleneck (Luikart and Cornuet 1998) as rare alleles, which contribute little to heterozygosity (Hedrick et al.1986), are removed quickly from the population through drift. Restoration of allelic diversity to pre-bottleneck levels would be accomplished through mutation (Stefenon et al. 2008) and/or migration, and might be anticipated to take a longer period of time than recovery of heterozygosity to pre-bottleneck levels. However, the discrepancy between the latitudinal breakpoints for heterozygosity and allelic richness was only slight (i.e. 2°). Sufficient time for recovery of shad genetic diversity to pre-bottleneck levels may have elapsed since post-glacial colonization of the Hudson, Connecticut, and Merrimack Rivers, and the possibility that this discrepancy has resulted from noise in the data cannot be ruled out.

4.4.2 Clinal Variation in Allele Frequencies

The significant clinal variation observed for the majority (10/13) of loci examined is consistent with previous genetic studies of shad (Shoubridge 1978; Nolan et al. 1991), and suggests that clinal patterns are a common feature of genetic variation in shad. Clinal patterns of genetic variation have been observed among other species with broad latitudinal ranges (e.g. *Fundulus heteroclitus*; Adams et al. 2006). Although selection is often invoked as a cause of clinal neutral genetic variation (e.g. Eanes 1999; Baines et al. 2004), clines can also arise as a consequence of founder events, and spatially restricted gene flow (Vasemägi 2006). No evidence of selection was detected for the microsatellite loci employed here, implying that the clinal variation detected is more likely due to spatial genetic structure, perhaps in combination with population founder events. The most common allele for each of Asa16, AsaD042, and AsaD429 exhibited a marked

increase in frequency among populations from formerly glaciated regions (Appendix 7; panels d, l, and m, respectively). These alleles may have 'surfed' to high densities on the wave of range expansion through drift during post-glacial colonization (Excoffier and Ray 2008).

4.4.3 RANGEWIDE POPULATION STRUCTURE

This study generally supports the hypothesis that individual rivers support genetically distinguishable shad spawning populations (Bentzen et al. 1989) whose genetic compositions are temporally stable over at least 2-3 year time scales (AMOVA). However, there were some cases in which genetic differentiation between particular U.S. spawning runs was not significant. These cases might be artefacts of small sample sizes (Nanticoke and Delaware Rivers), or indicative of high levels of gene flow (natural or human-mediated) in areas such as Chesapeake Bay and the region occupied by semelparous spawning runs. Despite weak genetic differentiation among U.S. spawning runs, a highly significant IBD pattern ($p<0.001$; $r=0.71$) was observed within the U.S. portion of the species range. Previous surveys of shad microsatellite variation also revealed weak population structure among U.S. spawning runs (Brown et al. 2000), but failed to detect significant IBD (Waters et al. 2000). Population structure among Canadian spawning runs was previously discussed (Chapter 3), and is not recapitulated here.

Bayesian based estimates of nine genetically distinguishable shad population groups among the 33 spawning runs assessed are likely conservative; these methods may exhibit limited power under scenarios involving low levels of genetic differentiation (Faubet et al. 2007). For example, although genic heterogeneity was detected among

most (41/45) comparisons south of Chesapeake Bay, pairwise genetic differentiation was weak (F_{ST}=0.001-0.019), and these spawning runs were grouped into the 'semelparous U.S. cluster' using Bayesian analyses. Further, AMOVA revealed a highly significant component of variation among rivers within clusters, supporting the notion that these Bayesian estimates of population structure are conservative.

Although the ΔK criterion supported the existence of only two genetic clusters across the species range, agreement of $\ln Pr(X|K)$ and ΔK based estimates have only been observed using data simulated under simplistic scenarios (i.e. island model of migration among four populations; Waples and Gaggiotti 2006). Results from this study agree with previous empirical studies that observed non-congruence between these estimators (Piñeiro et al. 2007; Vaha et al. 2007), and suggests that the ΔK statistic may have only detected 'deep-rooted' structure among spawning runs from across the species range (Piñeiro et al. 2007).

The discontinuities in gene flow identified among spawning runs from across the species range using BARRIER were identical to those revealed among Canadian spawning runs (Chapter 3). The lack of statistical support (i.e. absence of strongly supported consensus barriers) for discontinuities in gene flow among U.S. spawning runs likely reflects weak levels of genetic differentiation and a high degree of allele frequency overlap across the U.S. portion of the species range.

4.4.4 Regional Patterns of Isolation By Distance

This study detected different spatial patterns of shad population structure across the species' range. While significant IBD patterns were observed among iteroparous and

semelparous spawning runs, ANCOVA revealed heterogeneity of IBD slopes between these alternate reproductive strategies. However, these different IBD patterns were largely influenced by the inclusion of Canadian spawning runs. Spatial patterns of population structure between U.S. iteroparous and semelparous spawning runs were not significantly different, and it is unlikely that alternate reproductive strategies are responsible for the different patterns observed across the species' range.

It is more likely that these dissimilar patterns of population structure are due to different shad management strategies between the U.S. and Canada, and/or alternate glacial histories. My results suggest that stock transfers have had a negative effect on shad population structure in the U.S. (i.e. reduced genetic differentiation/genetic swamping; Bouzat et al. 2009), and ANCOVA revealed significantly different IBD patterns between Canadian and U.S. spawning runs. Inspection of residuals from linear regression of the IBD pattern among U.S. spawning runs revealed that the degree of differentiation among pairs of spawning runs subjected to stock transfers was lower than what might be predicted based on geographic distance alone, and that this effect increases with distance between source stock and the recipient spawning population (Figure 4.7b). However, significantly different IBD patterns were also observed among formerly glaciated and non-glaciated rivers, and the categorization of spawning runs into these groups broadly overlapped their classification as Canadian and U.S., respectively. Therefore, it remains uncertain whether differential patterns of population structure detected across the species range reflect the influence of different shad management strategies between the U.S. and Canada, effects of alternate glacial histories, or combinations of these factors.

4.4.5 FACTORS INFLUENCING U.S. POPULATION STRUCTURE

Although determining the relative influences of historical demography, microevolutionary processes, and anthropogenic effects on patterns of population structure is difficult, some of the observation made in this study may provide insight when taken in the greater context of the body of literature available for shad.

Shad spawning runs are hypothesized to have been confined to the southeastern U.S. during the LGM (Bentzen et al. 1989). Results of this study suggest that post-glacial range expansion from this region probably followed a stepwise process, with genetic drift shaping the initial allele frequency distributions of founder populations (Ibrahim et al. 1996). Principal coordinates analysis revealed that the factor that explained the greatest proportion of variation in genetic distance among spawning runs from across the species range was strongly associated with latitude ($p<0.001$; $r^2=0.85$), and suggests that the spatial distribution of shad genetic variation is underpinned by a stepping stone model of population structure. This interpretation is supported by the significant IBD pattern observed across the species' range.

Although a review of shad stocking history is beyond the scope of this study, prior summaries (Leach 1925, Hendricks 2003) and ongoing research (Hasselman and Hendricks, *in prep.*) suggest that stock transfers have not occurred among spawning runs south of Cape Fear River, NC. Thus, the weak genetic differentiation observed among semelparous spawning runs may have resulted from natural processes. Gene flow among semelparous spawning runs may be greater because of relaxed natal homing associated with the environmental stability hypothesis (Glebe and Leggett 1981). Temporal variation in river flow and temperature coincident with the period of early larval

development reduces the survival of shad larvae, and are the ultimate regulators of year class strength and recruitment to the spawning stock (Marcy 1976; Leggett 1977; Shoubridge and Leggett 1978; Crecco and Savoy 1987). Northern rivers exhibit greater temporal variation in these environmental variables during this critical period and are unpredictable, whereas southern rivers provide temporally stable and predictable environments for larval rearing (Shoubridge and Leggett 1978; Glebe and Leggett 1981). Although this rationale is primarily used to support the existence of shad semelparity, it may be just as valid a hypothesis to explain weak genetic structure among southern U.S. rivers. Semelparous shad may not experience reduced reproductive success if they do not home to their natal river with the same fidelity as their iteroparous counterparts, but still spawn among rivers within the semelparous portion of the species' range.

An alternative hypothesis for the weak genetic differentiation among semelparous spawning runs concerns the degree of overlap in the duration of the shad spawning season. The timing and duration of shad spawning is cued to thermal preferences (Leggett and Whitney 1972), and proceeds northward from the southern part of the species' range in a temporally progressive manner (Limburg et al. 2003). The duration of the 'temperature window' suitable for spawning (i.e. 14-18.5°C) varies from two to three months (November-February) for semelparous populations, to approximately three weeks (May-June) for iteroparous populations (Leggett and Whitney 1972). While the shorter duration of suitable spawning temperatures in the north may reduce the potential for interpopulation straying (Bentzen et al. 1989), the lengthier spawning season in the south may increase the potential for gene flow among semelparous spawning runs.

Shad stocking activities have historically focused on supplementation of iteroparous spawning runs (i.e. Chesapeake Bay, Delaware, Hudson, Connecticut, and New England rivers) (Hendricks 2003). Results from this study suggest that stock transfers have served to reduce the level of genetic differentiation (i.e. genetic swamping; Bouzat et al. 2009) among iteroparous spawning runs, and that this effect increases with distance between the donor and recipient spawning stock. Thus, the non-significant genetic differentiation detected among the majority of Chesapeake Bay spawning runs, and weak population structure observed among U.S. iteroparous stocks may reflect the influence of these stocking activities. Since 1971, the Susquehanna spawning run at the inner end of Chesapeake Bay has been supplemented with over 449 million shad larvae from several east coast rivers, including the Hudson and Delaware (Hendricks 2003, Brown and St. Pierre 2001). Between 1980 and 1987, 33,000 pre-spawn adults from the Hudson and Connecticut Rivers were transplanted to the Susquehanna (Hendricks 2003), with the Connecticut also serving as a brood source for several New England rivers, including the Merrimack. These efforts have been deemed 'successful' as hatchery shad constituted 46% of the 163,000 adult shad returning to the Susquehanna River in 2000 (Hendricks 2003).

Although stock transfers have likely influenced shad population structure in Chesapeake Bay, the weak genetic differentiation observed among Chesapeake Bay spawning runs may partially reflect an element of post-glacial colonization. Present day Chesapeake Bay represents a 'drowned' Susquehanna River valley that has been repeatedly altered by fluctuating sea levels (Hocutt et al. 1986). During the Wisconsinan glaciation all present day Chesapeake Bay drainages south to the James River were

tributaries of the 'Greater Susquehanna River' (Hocutt et al. 1986). Shad exhibit no evidence of population sub-structure among tributaries within four Canadian drainages (Chapter 3), and it is reasonable to postulate that initial genetic differentiation among tributaries of the post-glacially colonized 'Greater Susquehanna River' was weak or non-existent. The significant genic heterogeneity observed among most contemporary Chesapeake Bay spawning runs may reflect the influence of drift on allele frequencies since post-glacial colonization, and the weak population structure detected may reflect the relatively recent effects of stocking activities.

4.4.6 Contributions to Genetic Diversity

Shad exhibited increased within population genetic differentiation and decreased genetic diversity at the northern and southern edges of the species range relative to the central portion of their distribution, consistent with the central-marginal hypothesis (reviewed in Eckert et al. 2008). The tendency for genetic differentiation of peripheral populations from the core of species ranges is well documented (Lammi et al. 1999; Rossum et al. 2003; Hampe and Petit 2005; Hamilton and Eckert 2007), and may be due to genetic drift in populations with limited gene flow from the core of the species distribution, and to selection pressures that favor different genotypes at the core versus the edge of the range (Hoffman and Blows 1994; Kirkpatrick and Barton 1997; Hampe and Petit 2005).

Reduced levels of within population genetic diversity at the range margins begs the question: what is the conservation value of peripheral shad populations? While stochastic reductions of genetic diversity within peripheral populations may limit their current evolutionary potential (Hoffman and Blows 1994; Hoffman and Parsons 1997;

Gaston 2003; Blows and Hoffman 2005), preserving genetic diversity at range margins may be important for future adaptive potential (Gibson et al. 2009). In fact, rapid evolutionary change may be more likely at the range periphery because of the combined effects of genetic isolation from central populations, and different selection pressures (Thomas 2005).

Shad spawning runs from the United States made a greater contribution to global genetic diversity than Canadian stocks, and exhibited a greater amount of within population diversity for both observed heterozygosity and allelic richness, consistent with the spatial distribution of shad genetic variation (Figure 4.2). In the context of genetic conservation, allelic richness has been acknowledged as a more relevant measure of diversity than heterozygosity (Schoen and Brown 1993, El Mousadik and Petit 1996), as the preservation of alleles may be more important for population persistence than the maintenance of allele frequencies (Marshall and Brown 1975). While heterozygosity is proportional to the amount of genetic variance at a locus and is related to the immediate response to selection, the limit of this response (over several generations) is determined by the initial allelic composition of a population, regardless of the allelic frequencies (Allendorf 1986, James 1971). The lesser degree of among population allelic differentiation observed for U.S. spawning runs provides direct evidence that they share a high proportion of alleles, and is consistent with the distribution of alleles observed among US populations for multiple loci (Appendix 6).

4.5 CONCLUSIONS

Although a stepping stone model of population structure serves as the foundation for the magnitude and spatial distribution of shad neutral genetic variation from across the species' range, the relative influences of historical demography, microevolutionary processes, and anthropogenic factors on contemporary patterns of population structure vary depending on the portion of the range considered.

With the exception of Merrimack River, spawning runs included in this study that occupy formerly glaciated habitats have not been the recipients of stock transfers, and the spatial distribution of genetic variation among spawning runs in formerly glaciated areas probably reflects the combined effects of founder events or other population bottlenecks experienced during post-glacial colonization, as well as genetic drift and gene flow since that time. Sequential latitudinal declines in genetic diversity north of the LGM are probably associated with stepwise post-glacial range expansion and successive population founder events. The clinal variation in allele frequencies I observed may have resulted from spatially restricted gene flow, although some alleles may have 'surfed' to high densities on the wave of range expansion through drift during post-glacial colonization.

Stock transfers appear to have reduced the level of genetic differentiation among some iteroparous spawning runs, and the weak population structure observed within Chesapeake Bay may at least partly reflect the influence of historical stocking activities on contemporary population structure. However, the degree of differentiation among Chesapeake Bay spawning runs may be naturally low as a result of post-glacial

colonization of the 'Greater Susquehanna River' (Hocutt et al. 1986). Present day Chesapeake Bay drainages were tributaries of this proto-Susquehanna River during the Wisconsinan glaciation, and research within the Canadian portion of the species range (Chapter 3) suggests that shad do not exhibit population substructure among tributaries with drainages.

Although semelparous spawning runs exhibit weak differentiation, this is not likely to be the result of stocking practices, as stock transfers have not been documented among spawning runs south of the Cape Fear River, NC. Low levels of genetic differentiation among semelparous spawning runs may reflect increased levels of migration and gene flow associated with the environmental stability of southern rivers, or the greater duration of the spawning period in this portion of the species' range, or a combination of the two factors. The shorter duration of the spawning period in the northern portion of the range may reduce overlap in spawning times among rivers, thereby accentuating the effect of isolation by distance by adding the influence of an 'isolation by time' effect to the magnitude of genetic differentiation. The different spatial patterns of population structure observed across the species range have probably not resulted from alternate reproductive strategies. However, it is uncertain whether these patterns reflect the influence of different management strategies between the U.S. (stocking) and Canada (no stocking), remnant signals of alternative glacial histories, or combinations thereof.

Patterns of within population diversity and differentiation are consistent with the central-marginal hypothesis (reviewed in Eckert et al. 2008), and reduced genetic diversity at the northern and southern range margins may be a consequence of genetic

drift in populations with limited gene flow, or differential selection pressures. While U.S. spawning runs made a greater contribution to global genetic diversity than Canadian stocks, the lesser degree of among population allelic differentiation observed among U.S. spawning runs suggests that they share a high proportion of alleles, possibly as a consequence of some combination of natural or human-mediated gene flow. Although I cannot comment on the distribution of quantitative genetic variation, or the presence of local adaptations, lower levels of genetic variation exhibited by peripheral spawning runs (St. Lawrence River, St. John's River) and most Canadian stocks may restrict their ability to mount variable responses to future perturbations, and limit their capacity for long term persistence.

Shad restoration will benefit from this study, as knowledge of the magnitude and distribution of shad neutral genetic variation, and an understanding of the relative contributions of spawning runs to species genetic variation, will aid fisheries managers in the prioritization of stocks for focusing conservation resources and management efforts. While the ultimate impacts of stocking activities on the genetic integrity of U.S. shad spawning populations requires further examination, a recent genetic assessment of the Susquehanna River spawning population suggests that the proportional contribution of stocking sources has had a great influence on the gene frequencies and genetic structure of the stock (Julian and Bartron 2006). Out of basin stock transfers may have limited the long term adaptive potential of the Susquehanna spawning stock (Julian and Bartron 2006), and may have negatively impacted the evolutionary potential (Frankham 1995) of some other iteroparous shad spawning runs. Future restoration efforts should avoid stock transfers which can reduce population fitness through the loss of local adaptations and the

breakdown of co-adapted gene complexes (outbreeding depression) (McClelland and Naish 2007); an unfortunate outcome which is ultimately counterproductive to restoration goals for shad populations and the long-term persistence of the species.

CHAPTER 5

CONCLUSIONS

5.1 SUMMARY AND DISCUSSION

Biologists have long recognized that genetic variation within populations is crucial for evolutionary processes, adaptive potential, and the long term persistence of species. Intraspecific genetic diversity has been shown to improve colonization success (Gamfeldt et al. 2005), increase resistance to pathogens (Pearman and Garner 2005), enhance population stability (Bjornstad and Hansen 1994; Doebeli and de Jong 1999; Agashe 2009), and population persistence (Newman and Pilson 1997; Vilas et al. 2006). However, populations are not necessarily equivalent in their amount of genetic variation, or in their responses to future environmental conditions. Therefore, information about the magnitude and spatial distribution of intraspecific genetic variation is integral to conservation planning, and preserving species evolutionary potential (Frankham 1995).

A meta-analysis of neutral genetic variation for Nearctic fishes (Chapter 2) revealed that latitude is an important determinant of the spatial distribution of genetic diversity among and within species. Although latitudinal declines of genetic variation among species are to some degree mirrored by spatial declines within species, inter- and intraspecific patterns may differ substantially. For example, while linear declines were observed among anadromous teleosts, declines of intraspecific variation for Atlantic anadromous fishes were non-linear. The linear decline among anadromous teleosts may be due to the inclusion of northern species (>60°N) with inherently low genetic diversity,

or the use of a single measure of central tendency to represent the spatial distribution of intraspecific genetic variation (i.e. mean study latitude) in interspecific comparisons, or a combination of the two factors. Latitudinal declines of genetic variation within species encourage a cautionary approach for interspecific comparisons and inferences of broad spatial patterns of genetic variation, especially when data for individual species are obtained from only a portion of their range. In fact, non-linear declines of intraspecific genetic variation for species with broad geographic ranges (i.e. Atlantic anadromous fishes) suggest that a thorough understanding of the spatial distribution of intraspecific genetic variation requires a range-wide approach. Latitudinal declines of intraspecific genetic diversity are consistent with the presumed influence of population bottleneck/founder events associated with recurrent Pleistocene glaciations on spatial patterns of genetic variation. Consistent evidence for the influence of the LGM on the spatial distribution of intraspecific genetic variation for Atlantic anadromous fishes suggests that the portion of the Atlantic coast north of the LGM may constitute an empirical example of a one-dimensional stepping stone model of post-glacial colonization.

An examination of the spatial partitioning of neutral genetic variation for American shad within the Canadian portion of the species' range (Chapter 3) revealed temporally stable genetic differentiation among all drainages, supporting the hypothesis that spawning runs support genetically distinguishable populations (Bentzen et al. 1989). Temporally stable population sub-structure was not observed among tributaries within drainages, and suggests that shad natal homing is less effective at spatial scales finer than the drainage level. However, Bayesian methods identified seven clusters of genetically

distinguishable groups among 12 spawning runs, and provides evidence for metapopulation structure (sensu Hanski 1999) on regional spatial scales (<500km). A significant pattern of isolation by distance was observed among all spawning runs, consistent with a stepping stone model of population structure. Although regional (i.e. Bay of Fundy vs. non-Bay of Fundy) patterns of isolation by distance were similar, a greater degree of differentiation was detected among Bay of Fundy spawning runs, regardless of the spatial scale of comparison. Counterclockwise migration explained a greater proportion of genetic variation among Bay of Fundy spawning runs than 'direct route' distances, suggesting that pre-spawning shad follow the same migratory path hypothesized for non-spawning and post-spawning shad in the Bay of Fundy (Dadswell 1983; 1987). This pattern may be associated with the hydrodynamic features of the Bay of Fundy (Bumpus and Lauzier 1965) and the seasonal distribution of zooplankton (Locke and Corey 1989), and are likely recent in origin (6300 ±1100 ybp; Amos 1978).

A survey of the magnitude and spatial distribution of neutral genetic variation across the native distribution of American shad (Chapter 4) revealed different spatial patterns of genetic diversity among spawning runs from non-glaciated and formerly glaciated portions of the species' range. Sequential reductions of intraspecific genetic variation with latitude were observed among spawning populations from formerly glaciated regions, consistent with a stepwise process of post-glacial range expansion, and successive population founder events. This process may have lead to the significant clinal variation observed for the majority of loci examined, as alleles may have 'surfed' to high densities on the wave of range expansion via genetic drift during post-glacial colonization (Excoffier and Ray 2008). Relatively weak genetic differentiation and

population structure was observed among U.S. spawning, and is a consequence of the high degree of allele frequency overlap observed within this portion of the species' range (Appendix 6). Despite weak genetic differentiation among U.S. spawning runs, isolation by distance was detected across the species' range. Different spatial patterns of population structure revealed across the species' range are not likely due to the influence of alternative reproductive strategies (iteroparity vs. semelparity), but are a possible consequence of different management strategies in Canada (no stocking) and the United States (stocking), alternative glacial histories, or combinations thereof. Reciprocal patterns of genetic diversity and differentiation detected across the species' range suggests that U.S spawning runs contribute more to diversity and less to differentiation than Canadian spawning runs, and has implications for future shad restoration efforts.

This thesis represents the first range-wide survey of neutral genetic variation for an anadromous clupeid, and greatly advances our understanding of the magnitude and spatial distribution of genetic diversity for American shad. This research identifies important aspects of shad biology (e.g. latitudinal declines of intraspecific diversity, metapopulation structure, isolation by distance, different spatial patterns of population structure) that eluded previous surveys of shad genetic variation based on mtDNA (reviewed in Nolan et al. 2003) and microsatellites (Brown et al. 2000; Waters et al. 2000). However, prior microsatellite based studies of shad were limited in scope to few loci and few geographically proximal (<500 km) populations near the centre of the species distribution. This research highlights the importance of adequately sampling across a species distribution before making claims of range wide patterns of genetic variation.

Although historical population bottlenecks or founder events associated with recurrent Pleistocene glaciations have influenced the spatial distribution of shad neutral genetic variation, contemporary patterns also reflect the influences of microevolutionary processes since post-glacial colonization, and anthropogenic factors; the relative effects of which vary depending on the portion of the species range considered (see below).

The magnitude and spatial distribution of neutral genetic variation among Canadian spawning runs is more likely to reflect the influences of historical demography and microevolutionary processes than anthropogenic effects. The sequential reductions of intraspecific genetic variation with latitude observed in this portion of the species' range are a likely consequence of a stepwise process of post-glacial range expansion into Atlantic Canada, and successive population founder events. The affinity between Gulf of St. Lawrence spawning runs and the Saint John River revealed in Chapter 3 may reflect aspects of the species' historical biogeography in Atlantic Canada, and alterations to connectivity brought about by post-glacial sea level change. Although Canadian spawning runs have not been the subject of stock transfers, the influence of anthropogenic factors on Canadian spawning runs cannot be entirely discounted. Some Canadian spawning runs have been extirpated, and the effects of dams on the spatial distribution of shad genetic variation are uncertain. For instance, it remains unclear whether the exclusion of shad from spawning habitat above Mactaquac Dam in the Saint John River has resulted in the loss of a genetically distinguishable sub-population. Adult spawners heading further upstream but encountering the dam may have spawned in various tributaries below Mactaquac Dam, resulting in introgression which may have yet to reach equilibrium. Although the influence of stocking activities on patterns of shad

genetic variation are anticipated to be greater among U.S. spawning runs, the potential for indirect effects on Canadian stocks via straying and introgression from U.S. spawning runs supplemented with out of basin stock transfers, is unknown.

The relatively weak genetic differentiation and population structure observed among U.S. spawning runs is a consequence of the high degree of allele frequency overlap observed within this portion of the species' range. Although this research generally supports the hypothesis that shad spawning runs constitute genetically distinguishable spawning populations (Bentzen et al. 1989), rare instances of genic homogeneity and weak genetic differentiation among semelparous and Chesapeake Bay spawning runs suggest elevated levels of gene flow in these regions. Weak genetic differentiation among semelparous spawning runs are unlikely to reflect the influence of human-mediated gene flow, as stock transfers have not been documented among spawning runs south of the Cape Fear River, NC (agency completion reports; U.S. Fish Commission internal documents; Leach 1925; Hendricks 2003). Alternatively, weak population structure in this portion of the species' range may be a consequence of natural gene flow resulting from relaxed natal homing due to the relative stability of these rivers during the spawning season for environmental variables important for shad reproductive success (Glebe and Leggett 1981), or the comparatively lengthier spawning season (2-3 months; Leggett and Whitney 1972) observed in this portion of the species' range. These are not necessarily mutually exclusive hypotheses, as the duration of the spawning season is likely correlated with the environmental stability of these rivers. In this context, spawning run duration and relaxed natal homing may constitute proximate factors for

explaining weak genetic differentiation among semelparous shad, while environmental stability may represent the ultimate factor.

Weak genetic differentiation and population structure observed among Chesapeake Bay spawning runs may be a consequence of stock transfers, which appear to reduce the level of genetic differentiation between donor and recipient spawning population (i.e. genetic swamping; Bouzat et al. 2009). Iteroparous spawning runs have been the focus of supportive breeding and stock transfers for over a century, with much of the activity focused on restoration of Chesapeake Bay spawning populations (Hendricks 2003). However, genetic differentiation among Chesapeake Bay spawning runs may be naturally low due to post-glacial colonization of the 'Greater Susquehanna River' (Hocutt et al. 1986) when present day Chesapeake Bay drainages were tributaries of that river. The significant genic differentiation detected among some Chesapeake Bay spawning populations may reflect the influence of genetic drift on allele frequencies since post-glacial colonization, but it is uncertain whether genetic differentiation among these spawning runs was substantially greater prior to commencement of stocking activities in this region. Although the ultimate impacts of stocking on the genetic integrity of U.S. shad spawning runs remain unknown, a genetic assessment of the Susquehanna River spawning population suggests that the proportional contribution of stocking sources has influenced the gene frequencies and genetic structure of the stock (Julian and Bartron 2006). Out of basin stock transfers may have limited the long term adaptive potential of the Susquehanna spawning population (Julian and Bartron 2006), and may have negatively impacted the evolutionary potential (Frankham 1995) of other iteroparous shad spawning runs.

In the context of genetic conservation, the maintenance of allelic diversity has been suggested as more important for population persistence than the preservation of allele frequencies (Marshall and Brown 1975). While heterozygosity is related to the immediate response to selection, the limit of this response is determined by the initial allelic composition of a population, regardless of the allelic frequencies (James 1971; Allendorf 1986). The lesser degree of among population allelic differentiation observed for U.S. spawning runs (Figure 4.8b) suggests that they share a high proportion of alleles (consistent with the distribution of alleles observed among U.S. spawning runs for multiple loci; Appendix 6), and begs the question: to what extent are U.S. shad spawning runs 'evolutionarily exchangeable'? However, the use of neutral microsatellite markers as a proxy of genome wide nuclear variation has recently come under scrutiny (Väli et al. 2008; Ljungqvist et al. 2010), and patterns of neutral genetic variation revealed in this thesis cannot attest to the presence or influence of local adaptations, or impacts of intraspecific life history variation (i.e. degree of iteroparity), on the 'ecological exchangeability' of U.S. shad populations.

Anadromous clupeids serve as a major source of marine-derived nutrients far upstream in riverine habitat via excretion, gamete release, and decomposition of post-reproductive mortalities, and provide an important annual subsidy to the energy and nutrient budgets of these ecosystems (Garman and Macko 1998; MacAvoy et al. 2000). A spawning run of 1,000,000 semelparous shad in a southern river (such as the size of historic spawning runs) with average weight of 1.5kg/fish would release 180 metric tons of marine-derived nitrogen upon their death (based on 12% nitrogen content of whole fish) (Limburg et al. 2003). If shad semelparity has a heritable basis, then the stocking of

semelparous spawning runs with iteroparous stock (or vice versa) may have an effect on ecosystems and food webs from the bottom up.

5.2 Implications for Management

The high degree of allele frequency overlap detected among U.S. spawning runs has direct implications for species management, as the power to discriminate among populations using molecular techniques requires some extent of non-overlapping gene frequencies (Smouse and Chevillon 1998). The microsatellites examined in this study may not provide enough resolution to accurately (>95% correct) assign an American shad captured in mixed stock/coastal intercept fisheries to river of origin. The suite of microsatellite loci used in this thesis represents a random sample of molecular markers, and it is doubtful that a similar number of any other randomly chosen microsatellites could provide greater resolution than these loci. My research suggests that a substantially greater number of microsatellites may be required for accurate classification of shad to river of origin, but may quickly result in diminishing returns. Previous attempts to assign shad to river of origin using mtDNA observed low assignment success because of low resolution (Epifanio et al. 1995). Although accurate assignment to river of origin may be possible within the Canadian portion of the species' range, assignment of U.S. shad using microsatellites may only be possible to region of origin (i.e. Chesapeake Bay).

Although the detection of genic heterogeneity among Canadian spawning runs justifies the current 'river-level' approach to shad management, the detection of shad metapopulation structure suggests that fisheries managers in Canada need to be concerned with the loss of genetic variation from river to regional spatial scales. The

elevated level of stock structure observed within the Bay of Fundy may warrant particular consideration should development of tidal power proceed within the region. The detection of self-sustaining spawning activity and persistence of metapopulation structure among relatively small drainages in Canada warrants further consideration of historical accounts of shad from additional small drainages along the Atlantic coast of Nova Scotia and within the Bay of Fundy.

Substantial resources have been allocated to the restoration of shad populations (e.g. stocking, alteration to fish passage facilities, dam removal) with varying success (reviewed in Hendricks 2003; Cooke and Leach 2003; Olney et al. 2003; St. Pierre 2003; Weaver et al. 2003). Future restoration efforts should avoid stock transfers which can reduce population fitness through the loss of local adaptations and the breakdown of co-adapted gene complexes (outbreeding depression) (McClelland and Naish 2007), an unfortunate outcome which is ultimately counterproductive to shad restoration goals and the species' long-term persistence.

Greater success in the restoration of shad spawning runs may be achieved by providing access to historical spawning grounds, and will not jeopardize the genetic integrity of spawning stocks. Dams among major U.S. rivers are partly responsible for the decline in shad abundance across their native range (Bilkovic et al. 2002a), and have precluded access to approximately 4000 km of historical spawning grounds (Limburg et al. 2003). The reproductive success of shad has been associated with upstream spawning migration distance, and the greater availability of suitable prey and optimal water temperatures for growth (Limburg 1996). Thus, access to historical spawning grounds may provide for greater shad reproductive success, and the restoration of shad spawning

populations to self-sustaining levels. The decline of shad abundance across their range was not instantaneous, but occurred gradually over a period of decades. Fisheries managers (and the public) need to recognize that recovery of shad spawning runs by providing access to historical spawning grounds will similarly be a gradual process, and dam removal should not be expected to yield immediate dividends; no restoration measure will not constitute a 'quick fix'.

5.3 FUTURE DIRECTIONS

The spatial patterns of neutral genetic variation for American shad reported in this thesis provide the foundation for future investigations. Weak genetic differentiation among semelparous spawning runs in the absence of documented stock transfers suggest that migration and gene flow may be greater among these rivers than in other portions of the species' range. Shad spawning site fidelity (97%) was estimated from the Annapolis River, NS (Melvin et al. 1986) (the only spawning population examined in this thesis that did not exhibit evidence for significant admixture), but has been applied broadly across the species' range. No study has examined the potential for latitudinal variation in the degree of shad philopatry/spawning site fidelity. Estimates of migration rate (m) and the effective number of migrants ($N_e m$) between shad spawning populations would aid interpretation of the spatial patterns of genetic variation revealed in this research. These estimates could be assessed along with publicly available time series data (U.S. Geological Survey: National Water Information System) for environmental variables (water temperature and flow) important for shad reproductive success to examine the relevance of the environmental stability hypothesis (Glebe and Legget 1981) to contemporary patterns of population structure.

Although weak genetic differentiation and population structure among iteroparous U.S. spawning runs may have resulted from stocking activities, this has not been demonstrated unequivocally, and the impacts of stocking on shad population structure requires further examination. A crucial first step in this process involves understanding the history of stock transfers and supportive breeding since these practices began. A full accounting of the stocking history for American shad has not been compiled (records are available in agency completion reports and internal documents; Hendricks 2003), but would prove a valuable resource for fisheries managers, and the future conservation of shad genetic variation.

Although American shad have experienced dramatic declines in abundance across their native distribution, transplants of shad to U.S. Pacific coastal rivers have resulted in remarkable spread and increased abundance of the species in their introduced range. Shad were repeatedly introduced to the Sacramento (619,000 fry-Hudson River source) and Columbia Rivers (60,000 fry-Hudson River source; 850,000 fry-Susquehanna River source) in the late 1800s, and quickly dispersed and colonized additional drainages (Anon. 1895; Leach 1925). Shad rapidly expanded its distribution over 5000 km of Pacific coastline, and has been reported from Mexico (Hart 1973) to Russia (Petersen et al. 2003). Beyond its capacity for dispersal and colonization of novel habitat, invasive shad have also increased dramatically in abundance, and now constitute the single largest spawning run of anadromous fish in the Columbia River (>4 million fish annually) (Petersen et al. 2003). The ultimate success of the shad invasion may be founded in elevated propagule pressure (Ficetola et al. 2008), as recurrent stocking of shad served as repeated invasion events, likely increasing the probability of successful invasion (Drury

et al. 2007). Repeated stocking of large numbers of individuals from divergent source populations (i.e. Hudson and Susquehanna Rivers) may have brought together considerable amounts of genetic variation and novel genetic combinations, augmenting reductions in genetic diversity expected during successive founder events, thereby maintaining shad adaptive potential and inadvertently aiding the invasion of novel habitats (Dlugosch and Parker 2008).

The success of this introduction provides an opportunity to examine the evolutionary responses of American shad in a novel environment. Species introduced to novel environments can exhibit rapid evolutionary changes (Thompson 1998), and a life history variant appears to have arisen in the Columbia River. Increasing reports of 'mini-shad' (too small to be adults, yet too large to be juveniles; B. Shields, Oregon State University, pers. comm. 2006) are consistent with the finding that Columbia River shad exhibit an evolutionary adaptation for increased juvenile growth rate (Rottiers et al. 1992). It has also been suggested that 'mini-shad' may be resident, a notion consistent with the presence of 'yearling' (*sensu* Limburg 1998) shad from the Hudson River; a source population for the Columbia River introduction.

REFERENCES

Adams, S.M., Lindmeier, J.B., and Duvernell, D.D. 2006. Microsatellite analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, *Fundulus heteroclitus*. Molecular Ecology 15: 1109-1123.

Agashe, D. 2009. The stabilizing effect of intraspecific genetic variation on population dynamics in novel and ancestral habitats. The American Naturalist 174(2): 255-267.

Allendorf, F.W. 1986. Genetic drift and the loss of alleles versus heterozygosity. Zoo Biology 5: 181-190.

Allendorf, F.W. and Seeb, L.W. 2000. Concordance of genetic divergence among sockeye salmon populations at allozyme, nuclear DNA, and mitochondrial DNA markers. Evolution 54: 640-651.

Allendorf, F.W. and Lundquist, L.L. 2003. Introduction: Population biology, evolution, and control of invasive species. Conservation Biology 17: 24-30.

Allendorf, F.W., Leary, R.F., Hitt, N.P., Knudsen, K.L., Lundquist, L.L., and Spruell, P. 2004. Intercrosses and the US endangered species act: should hybridized populations be included as Westslope cutthroat trout? Conservation Biology 18: 1203-1213.

Amos, C.L. 1978. Post-glacial evolution of Minas Basin, NS: A sedimentological interpretation. Journal of Sedimentary Petrology 48(3): 965-981.

Anonymous. 1895. The shad: history of experiments in California. Bulletin of the United States Fish Commission 15: 404-427.

Atlantic States Marine Fisheries Commission (ASMFC). 2007. American shad stock assessment report. Stock Assessment Report No. 07-01. Volume I, 238pp.

Avise, J.C., Bowen, B.W., Lamb, T., Meylan, A.B., and Bermingham, E. 1992. Mitochondrial DNA evolution at a turtle's pace: evidence for low genetic variability and reduced microevolutionary rate in the Testudines. Molecular Biology and Evolution 9(3): 457-473.

Baby, M.C., Bernatchez, L., and Dodson, J.J. 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 9: 61-68.

Bagley, M.J. and Gall, G.A.E. 1998. Mitochondrial and nuclear DNA sequence variability among populations of rainbow trout (*Oncorhynchus mykiss*). Molecular Ecology 7: 945-961.

Baines, J.F., Das, A., Mousset, S., and Stephan, W. 2004. The role of natural selection in worldwide populations of *Drosophila ananassae*. Genetics 168: 1987-1998.

Ballard, J.W.O. and Whitlock, M.C. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13: 729-744.

Banks, M.A., Rashbrook, V.K., Calavetta, M.J., Dean, C.A., and Hedgecock, D. 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of chinook salmon (*Oncorhynchus tshawytscha*) in California's central valley. Canadian Journal of Fisheries and Aquatic Sciences 57: 915-927.

Beacham, T.D., McIntosh, B., and MacConnachie, C. 2002. Microsatellite identification of individual sockeye salmon in Barkley sound, British Columbia. Journal of Fish Biology 61: 1021-1032.

Beacham, T.D., Supernault, K.J., Wetklo, M., Deagle, B., Labaree, K., Irvine, J.R., Candy, J.R., Miller, K.M., Nelson, R.J., and Withler, R.E. 2003. The geographic basis for population structure in Fraser River chinook salmon (*Oncorhynchus tshawytscha*). Fishery Bulletin 101: 229-242.

Beaumont, M.A. 1999. Detecting population expansion and decline using microsatellites. Genetics 153: 2013-2029.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., and Bonhomme, F. 2004. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. *Laboratoire Génome, Populations, Interactions, CNRS UMR 5000.* Université de Montpellier II; Montpellier.

Bentzen, P., Brown, G.C., and Leggett, W.C. 1989. Mitochondrial DNA polymorphism, population structure, and life history variation in American shad (*Alosa sapidissima*). Canadian Journal of Fisheries and Aquatic Sciences 46: 1446-1454.

Bentzen, P. and Paterson, I. G. 2005. Genetic analyses of freshwater and anadromous alewife (*Alosa pseudoharengus*) populations from the St. Croix River, Maine/New Brunswick. Final Report to Maine Rivers, 3 Wade Street, Augusta, ME 04330.

Bermingham, E. and Avise, J.C. 1986. Molecular zoogeography of freshwater fishes in the Southeastern United States. Genetics 113: 939-965.

Bermingham, E., Forbes, S.H., Friedland, K., and Pla, C. 1991. Discrimination between Atlantic salmon (*Salmo salar*) of North American and European origin using restriction analyses of Mitochondrial DNA. Canadian Journal of Fisheries and Aquatic Sciences 48: 884-893.

Bernatchez, L. and Dodson, J.J. 1990. Mitochondrial DNA variation among anadromous populations of cisco (*Coregonus artedi*) as revealed by restriction analysis. Canadian Journal of Fisheries and Aquatic Sciences 47: 533-543.

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

Bernatchez, L., Colombani, F., and Dodson, J.J. 1991. Phylogenetic relationships among the subfamily *coregoninae* as revealed by mitochondrial DNA restriction analysis. Journal of Fish Biology 39: 283-290.

Bernatchez, L., Guyomard, R., and Bonhomme, F. 1992. DNA sequence variation of the mitochondrial control region among geographically and morphologically remote European brown trout *Salmo trutta* populations. Molecular Ecology 1: 161-173.

Bernatchez, L. and 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. 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. and Wilson, C.C. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. Molecular Ecology 7: 431-452.

Bernatchez, L. and Wilson, C.C. 1999. Diadromy and genetic diversity in Nearctic and Palearctic fishes: a reply. Molecular Ecology 8: 529-530.

Bernatchez, L., Dempson, J.B., and Martin, S. 1998. Microsatellite gene diversity analysis in anadromous arctic char, *Salvelinus alpinus*, from Labrador, Canada. Canadian Journal of Fisheries and Aquatic Sciences 55: 1264-1272.

Bickham, J., Patton, J., and Gallaway, B. J. 1992. Genetic studies of the Arctic cisco (*Coregonus autumnalis*) in Alaska, multiple stocks. US Minerals Management Service No. OCS Report.

Bickham, J.W., Wood, C.C., and Patton, J.C. 1995. Biogeographic implications of Cytochrome *b* sequences and allozymes in sockeye (*Oncorhynchus nerka*). Journal of Heredity 86: 140-144.

Bilkovic, D.M., Hershner, C.H., and Olney, J.E. 2002a. Macroscale assessment of American shad spawning and nursery habitat in the Mattaponi and Pamunkey Rivers, Virginia. North American Journal of Fisheries Management 22: 1176-1192.

Bilkovic, D.M., Olney, J.E., and Hershner, C.H. 2002b. Spawning of American shad (*Alosa sapidissima*) and striped bass (*Morone saxatilis*) in the Mattaponi and Pamunkey Rivers, Virginia. Fishery Bulletin 100: 632-640.

Bjornstad O.N. and Hansen T.F. 1994. Individual variation and population dynamics. Oikos 69: 169-171.

Blows, M.W., and Hoffman, A.A. 2005. A reassessment of limits to evolutionary change. Ecology 86: 1371-1384.

Bohonak, A.J. 1999. Dispersal, gene flow, and population structure. Quarterly Review of Biology 74: 21-45.

Bonhomme, M., Blancher, A., Cuartero, S., Chikhi, L., and Crouau-Roy, B. 2008. Origin and number of founders in an introduced insular primate: estimation from nuclear genetic data. Molecular Ecology 17: 1009-1019.

Bos, D.H., Gopurenko, D., Williams, R.N., and DeWoody, J.A. 2008. Inferring population history and demography using microsatellites, mitochondrial DNA, and major histocompatibility complex (MHC) genes. Evolution 62: 1458-1468.

Bouzat, J.L., Johnson, J.A., Toepfer, J.E., Simpson, S.A., Esker, T.L., and Westermeier, R.L. 2009. Beyond the beneficial effects of translocations as an effective tool for the genetic restoration of isolated populations. Conservation Genetics 10: 191-201.

Bowen, B.W. and Avise, J.C. 1990. Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: influence of zoogeographic factors and life-history patterns. Marine Biology 107: 371-381.

Bradbury, I.R., Coulson, M.W., Campana, S.E., and 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 (Suppl. C): 95-114.

Bradbury, I.R., Campana, S.E., and Bentzen, P. 2008. Low genetic connectivity in an estuarine fish with pelagic larvae. Canadian Journal of Fisheries and Aquatic Sciences 65:147-158.

Brown, B.L., Epifanio, J.M., Smouse, P.E., and Kobak, C.J. 1996. Temporal stability of mtDNA haplotype frequencies in American shad stocks: To pool or not to pool across years? Canadian Journal of Fisheries and Aquatic Sciences 53: 2274-2283.

Brown, B.L., Gunter, T.P., Waters, J.M., and Epifanio, J.M. 2000. Evaluating genetic diversity associated with propagation-assisted restoration of American shad. Conservation Biology 14: 294-303.

Brown, J.J. and St. Pierre, R. A. 2001. Restoration of American shad *Alosa sapidissima* populations in the Susquehanna and Delaware rivers, USA. Proceedings of the Marine Technology Society Symposium. Honolulu, Hawaii. November 6-8, 2001.

Brown, J.R., Beckenbach, A.T., and Smith, M.J. 1992. Influence of Pleistocene glaciations and human intervention upon mitochondrial DNA diversity in white sturgeon (*Acipenser transmontanus*) populations. Canadian Journal of Fisheries and Aquatic Sciences 49: 358-367.

Bulak, J.S., Thomason, C.S., Han, K.P., and Ely, B. 2004. Genetic variation and management of striped bass populations in the coastal rivers of South Carolina. North American Journal of Fisheries Management 24: 1322-1329.

Bumpus, D. F., and Lauzier, L. M. 1965. Surface circulation on the continental shelf off eastern North America between Newfoundland and Florida. American Geographical Society Serial Atlas of the Marine Environment 7. New York, New York.

Carscadden, J.E. and Leggett, W.C. 1975. Meristic differences in spawning population of American shad, *Alosa sapidissima*. Evidence for homing to tributaries in the St. John River, New Brunswick. Journal of the Fisheries Research Board of Canada 32: 653-660.

Cashon, R.E., Vanbeneden, R.J., and Powers, D.A. 1981. Biochemical genetics of *Fundulus heteroclitus* (L.) .IV. Spatial variation in gene frequencies of Idh-A, Idh-B, 6-Pgdh-A, and Est-S. Biochemical Genetics 19: 715-728.

Castric, V., Bonney, F., and Bernatchez, L. 2001. Landscape structure and hierarchical genetic diversity in the brook charr, *Salvelinus fontinalis*. Evolution 55: 1016-1028.

Cavalli-Sforza, L.L. and Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. American Journal of Human Genetics 19: 233-257.

Cavalli-Sforza, L.L., Menozzi, P., Piazza, A. 1994. The history and geography of human genes. Princeton University Press, Princeton.

Cena, C.J., Morgan, G.E., Malette, M.D., and Heath, D.D. 2006. Inbreeding, outbreeding and environmental effects on genetic diversity in 46 walleye (*Sander vitreus*) populations. Molecular Ecology 15: 303-320.

Chaput, G. and Bradford, R. G. 2003. American shad (*Alosa sapidissima*) in Atlantic Canada. Canadian Science Advisory Secretariat Research Document 2003/009.

Cooke, D.W. and Leach, S.D. 2003. Beneficial effects of increased river flow and upstream fish passage on anadromous alosine stocks. American Fisheries Society Symposium 35: 331-338.

Corander, J. and Marttinen, P. 2006. Bayesian identification of admixture events using multilocus molecular markers. Molecular Ecology 15: 2833-2843.

Corander, J., Marttinen, P., and Mantyniemi, S. 2006. A Bayesian method for identification of stock mixtures from molecular marker data. Fishery Bulletin 104: 550-558.

Crandall, K.A., Bininda-Emonds, O.R.P., Mace, G.M., and Wayne, R.K. 2000. Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution 15: 290-295.

Crecco,V.A. and Savoy,T. 1987. Review of recruitment mechanisms of the American shad: the critical period and match-mismatch hypotheses reexamined. American Fisheries Society Symposium 1: 455-468.

Cronin, M.A., Spearman, W.J., Wilmot, R.L., Patton, J.C., and Bickham, J.W. 1993. Mitochondrial DNA variation in chinook (*Oncorhynchus tshawytscha*) and chum salmon (*O. keta*) detected by restriction enzyme analysis of polymerase chain reaction (PCR) Products. Canadian Journal of Fisheries and Aquatic Sciences 50: 708-715.

Curry, R.A. 2007. Late glacial impacts on dispersal and colonization of Atlantic Canada and Maine by freshwater fishes. Quaternary Research 67: 225-233.

Dadswell, M.J., Melvin, G.D., and Williams P.J. 1983. Effects of turbidity and the temporal and spatial utilization of the inner Bay of Fundy by American shad (*Alosa sapidissima*) (Pisces: Clupeidae) and its relationship local fisheries. Canadian Journal of Fisheries and Aquatic Sciences 40 (Suppl. 1): 322-330.

Dadswell, M.J., Melvin, G.D., Williams, P.J., and Themelis, D.E. 1987. Influences of origin, life history, and chance on the Atlantic coast migration of American shad. American Fisheries Society Symposium 1: 313-330.

Danzmann, R.G., Ferguson, M.M., and Arndt, S.K.A. 1993. Mitochondrial DNA variability in Ontario and New York rainbow trout (*Oncorhynchus mykiss*). Canadian Journal of Zoology 71: 1923-1933.

Danzmann, R.G., Morgan, R.P., Jones, M.W., Bernatchez, L., and Ihssen, P.E. 1998. A major sextet of mitochondrial DNA phylogenetic assemblages extant in eastern North American brook trout (*Salvelinus fontinalis*): distribution and postglacial dispersal patterns. Canadian Journal of Zoology 76: 1300-1318.

Dawson, A.G. 1992. Ice age earth. Routledge Press, London. 293 pp.

DeHaan, P.W., Libants, S.V., Elliott, R.F., and Scribner, K.T. 2006. Genetic population structure of remnant lake sturgeon populations in the upper Great Lakes basin. Transactions of the American Fisheries Society 135: 1478-1492.

Deiner, K., Garza, J.C., Coey, R., and Girman, D.J. 2007. Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and man-made barriers in the Russian River, California. Conservation Genetics 8: 437-454.

Diaz, M., Leclerc, G.M., and Ely, B. 1997. Nuclear DNA markers reveal low levels of genetic divergence among Atlantic and Gulf of Mexico populations of striped bass. Transactions of the American Fisheries Society 126: 163-165.

Dlugosch, K.M. and Parker, I.M. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology **17**: 431-449.

Dodson, J.J. and Leggett, W.C. 1974. Behavior of adult American shad (*Alosa sapidissima*) homing to the Connecticut River from Long Island Sound. Journal of the Fisheries Research Board of Canada 30: 1847-1860.

Doebeli, M. and de Jong, G. 1999. Genetic variability in sensitivity to population density affects the dynamics of simple ecological models. Theoretical Population Biology 55: 37-52.

Dowling, T.E., Minckley, W.L., and Marsh, P.C. 1996. Mitochondrial DNA diversity within and among populations of razorback sucker (*Xyrauchen texanus*) as determined by restriction endonuclease analysis. Copeia 542-550.

Drury, K.L.S., Drake, J.M., Lodge, D.M., and Dwyer, G. 2007. Immigration events dispersed in space and time: factors affecting invasion success. Ecological Modeling **206**: 63-78.

Duvernell, D.D., Lindmeier, J.B., Faust, K.E., and Whitehead, A. 2008. Relative influences of historical and contemporary forces shaping the distribution of genetic variation in the Atlantic killifish, *Fundulus heteroclitus*. Molecular Ecology 17: 1344-1360.

Dyke, A.S. and Prest, V.K. 1987. Late Wisconsinian and Holocene history of the Laurentide ice sheet. Geographic Physique et Quartenaire 4: 237-263.

Eanes, W.F. 1999. Analysis of selection on enzyme polymorphisms. Annual Review of Ecology and Systematics 30: 301-326.

Eckert, C.G., Samis, K.E., and Lougheed, S.C. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Molecular Ecology 17: 1170-1188.

El Mousadik, A., and Petit, R.J. 1996. High level of genetic differentiation for allelic richness among populations of argan tree [*Argania spinosa* (L.) skeels] endemic to Morocco. Theoretical and Applied Genetics 92: 832-839.

Elphinstone, M.S., Hinten, G.N., Anderson, M.J., and Nock, C.J. 2003. An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. Molecular Ecology Notes 3: 317-320.

Epifanio, J.M., Smouse, P.E., Kobak, C.J., and Brown, B.L. 1995. Mitochondrial DNA divergence among populations of American shad (*Alosa sapidissima*): How much variation is enough for mixed-stock analysis? Canadian Journal of Fisheries and Aquatic Sciences 52: 1688-1702.

Epifanio, J.M., Koppelman, J.B., Nedbal, M.A., and Philipp, D.P. 1996. Geographic variation of paddlefish allozymes and mitochondrial DNA. Transactions of the American Fisheries Society 125: 546-561.

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

Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin v. 3.1: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1: 47-50.

Excoffier, L. and Ray, N. 2008. Surfing during population expansions promotes genetic revolutions and structuration. Trends in Ecology and Evolution 23: 347-351.

Falush, D., Stephens, M., and Pritchard, J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Faria, R., Wallner, B., Weiss, S., and Alexandrino, P. 2004. Isolation and characterization of eight dinucleotide microsatellite loci from two closely related clupeid species (*Alosa alosa* and *A. fallax*). Molecular Ecology Notes 4: 586-588.

Faubet, P., Waples, R.S., and Gaggiotti, O.E. 2007. Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. Molecular Ecology 16: 1149-1166.

Fave, M.J. and Turgeon, J. 2008. Patterns of genetic diversity in Great Lakes bloaters (*Coregonus hoyi*) with a view to future reintroduction in Lake Ontario. Conservation Genetics 9: 281-293.

Ferguson, M.M., Bernatchez, L., Gatt, M.H., Konkle, B.R., Lee, S., Malott, M.L., and McKinley, R.S. 1993. Distribution of mitochondrial DNA variation in lake sturgeon (*Acipenser fulvescens*) from the Moose River basin, Ontario, Canada. Journal of Fish Biology 43: 91-101.

Ferguson, M.M. and Danzmann, R.G. 1998. Role of genetic markers in fisheries and aquaculture: useful tools or stamp collecting? Canadian Journal of Fisheries and Aquatic Sciences 55: 1553-1563.

Ficetola, G.F., Bonin, A., and Miaud, C. 2008. Population genetic reveals origin and number of founders in a biological invasion. Molecular Ecology 17(3): 773-782.

Fox, M.G., Claussen, J.E., and Philipp, D.P. 1997. Geographic patterns in genetic and life history variation in pumpkinseed populations from four east-central Ontario watersheds. North American Journal of Fisheries Management 17: 543-556.

Frankham, R. 1995. Conservation genetics. Annual Review of Genetics 29: 305-327.

Fraser, D.J. and Bernatchez, L. 2005. Allopatric origins of sympatric brook charr populations: colonization history and admixture. Molecular Ecology 14: 1497-1509.

Fulton, R.J. and Andrews J.T. 1987. La calotte glaciaire laurentidienne. Géographie Physique et Quaternaire 41.

Gach, M.H. 1996. Geographic variation in mitochondrial DNA and biogeography of *Culaea inconstans* (*Gasterosteidae*). Copeia 3:563-575.

Gagnon, M.C. and Angers, B. 2006. The determinant role of temporary proglacial drainages on the genetic structure of fishes. Molecular Ecology 15: 1051-1065.

Gamfeldt, L., Wallen, J., Jonsson, P.R., Berntsson, K.M., and Havenhand J.N. 2005. Increasing intraspecific diversity enhances settling success in marine invertebrates. Ecology 86: 3219-3224.

Gaston, K.J. 2003. The structure and dynamics of geographic ranges. Oxford University Press, Oxford, UK.

Garman, G.C. and Macko, S.A. 1998. Contribution of marine-derived organic matter to an Atlantic coast, freshwater, tidal stream by anadromous Clupeid fishes. Journal of the North American Benthological Society 17(3): 277-285.

George, A.L., Caldieraro, J.B., Chartrand, K.M., and Mayden, R.L. 2008. Population genetics of the blue shiner, *Cyprinella caerulea*. Southeastern Naturalist 7:637-650.

Gibson, S.Y., Van Der Marel, R.C., Starzomski, B.M. 2009. Climate change and conservation of leading edge peripheral populations. Conservation Biology 23(6): 1369-1373.

Girard, P. and Angers, B. 2006. The impact of postglacial marine invasions on the genetic diversity of an obligate freshwater fish, the longnose dace (*Rhinichthys cataractae*), on the Quebec peninsula. Canadian Journal of Fisheries and Aquatic Sciences 63: 1429-1438.

Giuffra, E., Bernatchez, L., and Guyomard, R. 1994. Mitochondrial control region and protein coding genes sequence variation among phenotypic forms of brown trout *Salmo trutta* from northern Italy. Molecular Ecology 3: 161-171.

Glebe, B.D. and Leggett, W.C. 1981. Latitudinal differences in energy allocation and use during the freshwater migrations of American shad (*Alosa sapidissima*) and their life history consequences. Canadian Journal of Fisheries and Aquatic Science 38: 806-820.

Gonzalez-Villasenor, L.I. and Powers, D.A. 1990. Mitochondrial DNA restriction site polymorphisms in the teleost *Fundulus heteroclitus* support secondary intergradation. Evolution 44: 27-37.

Goudet, J. 1995. FSTAT v. 1.2. A computer program to calculate *F*-statistics. Journal of Heredity 86: 485–486.

Goudet, J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9.3. Available at: http://www.unil.ch/izea/softwares/fstat.html.

Grunwald, C., Stabile, J., Waldman, J.R., Gross, R., and Wirgin, I. 2002. Population genetics of shortnose sturgeon *Acipenser brevirostrum* based on mitochondrial DNA control region sequences. Molecular Ecology 11: 1885-1898.

Grunwald, C., Maceda, L., Waldman, J., Stabile, J., and Wirgin, I. 2008. Conservation of Atlantic sturgeon *Acipenser oxyrinchus oxyrinchus*: delineation of stock structure and distinct population segments. Conservation Genetics 9: 1111-1124.

Guenette, S., Fortin, R., and Rassart, E. 1993. Mitochondrial DNA variation in lake sturgeon (*Acipenser fulvescens*) from the St. Lawrence River and James Bay drainage basins in Quebec, Canada. Canadian Journal of Fisheries and Aquatic Sciences 50: 659-664.

Gyllensten, U. 1985. The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous, and fresh water species. Journal of Fish Biology 26: 691-699.

Habicht, C., Seeb, L.W., and Seeb, J.E. 2007. Genetic and ecological divergence defines population structure of sockeye salmon populations returning to Bristol Bay, Alaska, and provides a tool for admixture analysis. Transactions of the American Fisheries Society 136: 82-94.

Hamilton, J.A., and Eckert C.G. 2007. Population genetic consequences of geographic disjunction: a prairie plant isolated on Great Lakes alvars. Molecular Ecology 16: 1647-1660.

Hampe, A., and Petit, R.J. 2005. Conserving biodiversity under climate change: the rear edge matters. Ecology letters 8: 461-467.

Hanski, I. 1999. Metapopulation Ecology. Oxford University Press, New York.

Hart, J.L. 1973. Pacific fishes of Canada. Fisheries Research Board of Canada Bulletin 180.

Hedrick, P.W., Brussard, P.F., Allendorf, F.W., Beardmore, J.A., and Orzack, S. 1986. Protein variation, fitness, and captive propagation. Zoo Biology 5; 91-99.

Hedrick, P.W. 2005. A standardized genetic differentiation measure. Evolution 59: 1633-1638.

Heist, E.J. and Mustapha, A. 2008. Rangewide genetic structure in paddlefish inferred from DNA microsatellite loci. Transactions of the American Fisheries Society 137: 909-915.

Hendricks, M.L., Hoopes, R.L., Arnold, D.A., and Kaufmann, M.L. 2002. Homing of hatchery-reared American shad to the Lehigh River, a tributary to the Delaware River. North American Journal of Fisheries Management 22: 243-248.

Hendricks, M.L. 2003. Culture and transplant of Alosines in North America. American Fisheries Society Symposium 35: 303-312.

Hewitt, G.M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society 58: 247-276.

Hewitt, G.M. 2000. The genetic legacy of the Quartenary ice ages. Nature 405: 907-913.

Hocutt, C.H., Jenkins, R.E., and Stauffer, J.R. Jr. 1986. Zoogeography of the fishes of the central Appalachians and central Atlantic coastal plain. *In* Zoogeography of North American Freshwater Fishes. Hocutt, C. and Wiley, E. O. eds, John Wiley and Sons. New York. Pg. 161-211.

Hoffman, A.A., Blows, M.W. 1994. Species borders; ecological and evolutionary perspectives. Trends in Ecology and Evolution 9:223-227.

Hoffman, A.A. and Parsons, P.A. 1997. Extreme environmental change and evolution. Cambridge University Press, Cambridge, UK.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6: 65-70.

Hubisz, M.J., Falush, D., Stephens, M., and Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9: 1322-1332.

Ibrahim, K.M., Nichols, R.A., and Hewitt, G.M. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity 77: 282-291.

Israel, J.A., Cordes, J.F., Blumberg, M.A., and May, B. 2004. Geographic patterns of genetic differentiation among collections of green sturgeon. North American Journal of Fisheries Management 24: 922-931.

James, J.W. 1971. The founder effect and response to artificial selection. Genetical Research 12: 249-266.

Jensen, J.L., Bohonak, A.J., and Kelley, S.T. 2005. Isolation by distance web service. BMC Genetics 6(1): 13.

Johnson, M.A. and Banks, M.A. 2008. Genetic structure, migration, and patterns of allelic richness among coho salmon (*Oncorhynchus kisutch*) populations of the Oregon coast. Canadian Journal of Fisheries and Aquatic Sciences 65: 1274-1285.

Julian, S.E. and Bartron, M.L. 2006. Genetics assessment of Susquehanna River American shad. *In* Restoration of American shad to the Susquehanna River-Annual Progress report. Susquehanna River Anadromous Fish Restoration Cooperative. pg. 216-241.

Julian, S.E. and Bartron, M.L. 2007. Microsatellite DNA markers for American shad (*Alosa sapidissima*) and cross-species amplification within the family Clupeidae. Molecular Ecology Notes 7: 805-807.

Kijima, A. and Matsunami, D. 1992. Haplotypic differences and variability of Mitochondrial DNA among cultured stocks of the Masu salmon complex. Nippon Suisan Gakkaishi 58: 1431-1436.

King, T.L., Spidle, A.P., Eackles, M.S., Lubinski, B.A., and Schill, W.B. 2000. Mitochondrial DNA diversity in North American and European Atlantic salmon with emphasis on the downeast rivers of Maine. Journal of Fish Biology 57: 614-630.

King, T.L., Lubinski, B.A., and Spidle, A.P. 2001. Microsatellite DNA variation in Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) and cross-species amplification in *Acipenseridae*. Conservation Genetics 2: 103-119.

Kirkpatrick, M., and N.H. Barton. 1997. Evolution of a species' range. The American Naturalist 150:1-23.

Krieger, J., and Fuerst, P.A. 2002. Evidence for a slowed rate of a molecular evolution in the order Acipenseriformes. Molecular Biology and Evolution 19(6): 891-897.

Lafontaine, P. and Dodson, J.J. 1997. Intraspecific genetic structure of white sucker (*Catostomus commersoni*) in northeastern North America as revealed by mitochondrial DNA polymorphism. Canadian Journal of Fisheries and Aquatic Sciences 54: 555-565.

Lammi, A., Siikamäki, P., and Mustajärvi, K. 1999. Genetic diversity, population size, and fitness in central and peripheral populations of a rare plant, *Lynchis viscaria*. Conservation Biology 12: 1069-1078.

Langella, O. 1999. Populations 1.2.30: a population genetic software. CNRS UPR9034 (Available at http://bioinformatics.org/~tryphon/populations).

Latch, E.K., Dharmarajan, G., Glaubitz, J.C., and Rhodes, O.E. 2006. Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conservation Genetics 7: 295-302.

Laughlin, T.F. and Turner, B.J. 1996. Hypervariable DNA markers reveal high genetic variability within striped bass populations of the lower Chesapeake Bay. Transactions of the American Fisheries Society 125: 49-55.

Leach, G. C. 1925. Artificial propagation of shad. U.S. Bureau of Fisheries Document No. 981.

Leberg, P.L. 2002. Estimating allelic richness: effects of sample size and bottlenecks. Molecular Ecology 11: 2445-2449.

Leggett, W.C., and Whitney, R.R. 1972. Water temperatures and the migrations of American shad. Fishery Bulletin 70: 659-670.

Leggett, W.C. 1977a. Ocean migration rates of American shad (*Alosa sapidissima*). Journal of the Fisheries Research Board of Canada 34: 1422-1426.

Leggett, W.C. 1977b. Density dependence, density independence and recruitment in the American shad (*Alosa sapidissima*) population of the Connecticut River, p. 3-17. *In* W. VanWinkle [ed.] Assessing the effects of power plant induced mortality on fish populations. Pergamon Press, New York. 380 p.

Leggett, W.C. and Carscadden, J.E. 1978. Latitudinal variation in reproductive characteristics of American shad (*Alosa sapidissima*): evidence for population specific life history stages in fish. Journal of the Fisheries Research Board of Canada 35: 1469-1478.

Leim, A.H. 1924. The life history of the shad (*Alosa sapidissima*) with special reference to the factors limiting its abundance. Contributions to Canadian Biology 2: 163-284.

Limburg, K.E. 1996. Modelling the ecological constraints on growth and movement of juvenile American shad (*Alosa sapidissima*) in the Hudson River Estuary. Estuaries 19(4): 794-813.

Limburg, K.E. 1998. Anomalous migrations of anadromous herrings revealed with natural chemical tracers. Canadian Journal of Fisheries and Aquatic Sciences **55**: 431-437.

Limburg, K.E., Hattala, K.A., and Kahnle, A. 2003. American Shad in its native range. American Fisheries Society Symposium 35 125-140.

Lin, J., Ziegler, E., Quinn, T.P., and Hauser, L. 2008. Contrasting patterns of morphological and neutral genetic divergence among geographically proximate populations of sockeye salmon *Oncorhynchus nerka* in Lake Aleknagik, Alaska. Journal of Fish Biology 73: 1993-2004.

Ljungqvist, M., Åkenson, M., and Hansson, B. 2010. Do microsatellites reflect genome-wide genetic diversity in natural populations? A comment on Väli *et al.* (2008). Molecular Ecology 19: 851-855.

Locke, A. and Corey, S. 1989. Amphipods, isopods and surface currents: a case for passive dispersal in the Bay of Fundy, Canada. Journal of Plankton Research 11: 419-430.

Lu, G., Basley, D.J., and Bernatchez, L. 2001. Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. Molecular Ecology 10: 965-985.

Luikart, G. and Cornuet, J.M. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12: 228-237.

Lundrigan, T.A., Reist, J.D., and Ferguson, M.M. 2005. Microsatellite genetic variation within and among Arctic charr (*Salvelinus alpinus*) from aquaculture and natural populations in North America. Aquaculture 244: 63-75.

Lutz-Carrillo, D.J., Nice, C.C., Bonner, T.H., Forstner, M.R.J., and Fries, L.T. 2006. Admixture analysis of Florida largemouth bass and northern largemouth bass using microsatellite loci. Transactions of the American Fisheries Society 135: 779-791.

MacAvoy, S.E., Macko, S.A., McIninch, S.P., and Garman, G.C. 2000. Marine nutrient contributions to freshwater apex predators. Oecologia 122: 568-573.

Mace, G.M. and Purvis, A. 2008. Evolutionary biology and practical conservation: bridging a widening gap. Molecular Ecology 17: 9-19.

Manni, F., Guerard, E., and Heyer, E. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Human Biology 76: 173-190.

Marcy, B.C. Jr. 1976. Early life history studies of American shad in the lower Connecticut River and effects of the Connecticut Yankee Plant, p. 141-168. *In* D. Merriman and L.M. Thorpe [ed.] The Connecticut River Ecological Study. Monograph No. 1. American Fisheries Society, Washington, DC.

Marshall, D.R. and Brown, A.H.D. 1975. Optimum sampling strategies in genetic conservation. Pages 53-80 *in* O.H. Frankel and J.G. Hawkes, editors. Crop genetic resources for today and tomorrow. Cambridge University Press, Cambridge, UK.

Marshall, H.D., Coulson, M.W., and Carr, S.M. 2009. Near neutrality, rate heterogeneity, and linkage govern mitochondrial genome evolution in Atlantic cod (*Gadus morhua*) and other Gadine fish. Molecular Biology and Evolution 26: 579-589.

McClelland, E.K. and Naish, K.A. 2007. What is the fitness outcome of crossing unrelated fish populations? A meta-analysis and an evaluation of future research directions. Conservation Genetics 8: 397-416.

McConnell, S.K.J., Ruzzante, D.E., O'Reilly, P.T., Hamilton, L., and Wright, J.M. 1997. Microsatellite loci reveal highly significant genetic differentiation among Atlantic salmon (*Salmo salar* L.) stocks from the east coast of Canada. Molecular Ecology 6: 1075-1089.

McCusker, M.R., Parkinson, E., and Taylor, E.B. 2000. Mitochondrial DNA variation in rainbow trout (*Oncorhynchus mykiss*) across its native range: testing biogeographical hypotheses and their relevance to conservation. Molecular Ecology 9: 2089-2108.

McDowall, R.M. 1999. Diadromy and genetic diversity in Nearctic and Palearctic fishes. Molecular Ecology 8: 527-528.

McElroy, T.C., Kandl, K.L., Garcia, J., and Trexler, J.C. 2003. Extinction-colonization dynamics structure genetic variation of spotted sunfish (*Lepomis punctatus*) in the Florida everglades. Molecular Ecology 12: 355-368.

McParland, T.L., Ferguson, M.M., and Liskauskas, A.P. 1999. Genetic population structure and mixed-stock analysis of walleyes in the Lake Erie-Lake Huron corridor

using allozyme and mitochondrial DNA markers. Transactions of the American Fisheries Society 128: 1055-1067.

McPhail, J.D. and Lindsey, C.C. 1970. Freshwater fishes of northwestern Canada and Alaska. Fisheries Research Board of Canada, Bulletin No. 173.

McQuown, E., Krueger, C.C., Kincaid, H.L., Gall, G. A.E., and May, B. 2003. Genetic comparison of lake sturgeon populations: Differentiation based on allelic frequencies at seven microsatellite loci. Journal of Great Lakes Research 29: 3-13.

Meirmans, P.G. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. Evolution 60: 2399-2402.

Melvin, G.D., Dadswell, M.J., and Martin, J.D. 1986. Fidelity of American shad, *Alosa sapidissima* (Clupeidae), to its river of previous spawning. Canadian Journal of Fisheries and Aquatic Sciences 46: 640-646.

Melvin, G.D., Dadswell, M.J., and McKenzie, J.A. 1992. Usefulness of meristic and morphometric characters in discriminating populations of American shad (*Alosa sapidissima*) (Ostreichthyes, Clupeidae) inhabiting a marine-environment. Canadian Journal of Fisheries and Aquatic Sciences 49: 266-280.

Mulligan, T.J. and Chapman, R.W. 1989. Mitochondrial DNA analysis of Chesapeake Bay white perch, *Morone americana*. Copeia 3: 679-688.

Mulvey, M., Newman, M.C., Vogelbein, W.K., Unger, M.A., and Ownby, D.R. 2003. Genetic structure and mtDNA diversity of *Fundulus heteroclitus* populations from polycyclic aromatic hydrocarbon-contaminated sites. Environmental Toxicology and Chemistry 22: 671-677.

Murdoch, M.H. and Hebert, P.D.N. 1997. Mitochondrial DNA evidence of distinct glacial refugia for brown bullhead (*Ameiurus nebulosus*) in the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 54: 1450-1460.

Nedbal, M.A. and Philipp, D.P. 1994. Differentiation of mitochondrial DNA in largemouth bass. Transactions of the American Fisheries Society 123: 460-468.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences USA 70: 3321-3323.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583-590.

Nei, M. and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97: 145-163.

Nei, M., Tajima, F., and Tateno, Y. 1983. Accuracy of estimated phylogenetic trees from molecular data. Journal of Molecular Evolution 19(2): 153-170.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York. 512 pp.

Neves, R.J., and Depres, L. 1979. The oceanic migration of American shad, *Alosa sapidissima*, along the Atlantic coast. Fisheries Bulletin 77: 199-212.

Newman, D., and Pilson, D. 1997. Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. Evolution 51: 354-362.

Nolan, K., Grossfield, J., and Wirgin, I. 1991. Discrimination among Atlantic coast populations of American Shad (*Alosa sapidissima*) using mitochondrial DNA. Canadian Journal of Fisheries and Aquatic Sciences 48: 1724-1734.

Nolan, K.A., Waldman, J.R., and Wirgin, I. 2003. Intraspecific and interspecific molecular differentiation of American shad and Alabama shad: a synthesis. American Fisheries Society Symposium 35: 295-302.

O' Reilly, P., Reimchen, T.E., Beech, R., and Strobeck, C. 1993. Mitochondrial DNA in *Gasterosteus* and Pleistocene glacial refugium on the Queen Charlotte Islands, British-Columbia. Evolution 47: 678-684.

Olney, J.E., Hopler, D.A. Jr., Gunter, T.P. Jr., Maki, K.L., and Hoenig, J.M. 2003. Signs of recovery of American shad in the James River, Virginia. American Fisheries Society Symposium 35: 323-330.

Olsen, J.B., Habicht, C., Reynolds, J., and Seeb, J.E. 2004. Moderately and highly polymorphic microsatellites provide discordant estimates of population divergence in sockeye salmon, *Oncorhynchus nerka*. Environmental Biology of Fishes 69: 261-273.

Orti, G., Bell, M.A., Reimchen, T.E., and Meyer, A. 1994. Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. Evolution 48: 608-622.

Paetkau, D., Calvert, W., Stirling, I., and Strobeck, C. 1995. Microsatellite Analysis of Population-Structure in Canadian Polar Bears. Molecular Ecology 4: 347-354.

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

Page, K.S., Scribner, K.T., and Burnham-Curtis, M. 2004. Genetic diversity of wild and hatchery lake trout populations: relevance for management and restoration in the Great Lakes. Transactions of the American Fisheries Society 133: 674-691.

Parker, K.M., Sheffer, R.J., and Hedrick, P.W. 1999. Molecular variation and evolutionarily significant units in the endangered Gila topminnow. Conservation Biology 13: 108-116.

Patton, J.C., Gallaway, B.J., Fechhelm, R.G., and Cronin, M.A. 1997. Genetic variation of microsatellite and mitochondrial DNA markers in broad whitefish (*Coregonus nasus*) in the Colville and Sagavanirktok rivers in northern Alaska. Canadian Journal of Fisheries and Aquatic Sciences 54: 1548-1556.

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

Pearman, P.B. and Garner, T.W.J. 2005. Susceptibility of Italian agile frog populations to an emerging strain of *Ranavirus* parallels population genetic diversity. Ecology Letters 8: 401-408.

Pearse, D.E., Donohoe, C.J., and Garza, J.C. 2007. Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environmental Biology of Fishes 80: 377-387.

Petersen, J.H., Hinrichsen, R.A., Gadomski, D.M., Feil, D.H., and Rondorf, D.W. 2003. American shad in the Columbia River. American Fisheries Society Symposium **35**: 141-155.

Peterson, D.L., Schueller, P., DeVries, R., Fleming, J., Grunwald, C., and Wirgin, I. 2008. Annual run size and genetic characteristics of Atlantic sturgeon in the Altamaha River, Georgia. Transactions of the American Fisheries Society 137: 393-401.

Petit, R.J., El Mousadik, A., and Pons, O. 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12: 844-855.

Pfrender, M.E., Hicks, J., and Lynch, M. 2004. Biogeographic patterns and current distribution of molecular genetic variation among populations of speckled dace, *Rhinichthys osculus* (Girard). Molecular Phylogenetics and Evolution 30: 490-502.

Phelps, S.R., Leclair, L.L., Young, S., and Blankenship, H.L. 1994. Genetic diversity patterns of chum salmon in the Pacific northwest. Canadian Journal of Fisheries and Aquatic Sciences 51: 65-83.

Philipp, D.P., Childers, W.F., and Whitt, G.S. 1985. Correlations of allele frequencies with physical and environmental variables for populations of largemouth bass, *Micropterus salmoides* (Lacepede). Journal of Fish Biology 27: 347-365.

Pielou, E.C. 1991. The return of life to glaciated North America. The University of Chicago Press, Chicago.

Piñeiro, R., Aguilar, J.F., Munt, D.D., and Feliner, G.N. 2007. Ecology matters: Atlantic-Mediterranean disjunction in the sand-dune shrub *Armeria pungens* (Plumbaginaceae). Molecular Ecology 16: 2155-2171.

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

Quattro, J.M., Leberg, P.L., Douglas, M.E., and Vrijenhoek, R.C. 1996. Molecular evidence for a unique evolutionary lineage of endangered Sonoran desert fish (Genus *Poeciliopsis*). Conservation Biology 10: 128-135.

Quattro, J.M., Greig, T.W., Coykendall, D.K., Bowen, B.W., and Baldwin, J.D. 2002. Genetic issues in aquatic species management: the shortnose sturgeon (*Acipenser brevirostrum*) in the southeastern United States. Conservation Genetics 3: 155-166.

Ramstad, K.M., Foote, C.J., and Olsen, J.B. 2003. Genetic and phenotypic evidence of reproductive isolation between seasonal runs of sockeye salmon in Bear Lake, Alaska. Transactions of the American Fisheries Society 132: 997-1013.

Ramstad, K.M., Woody, C.A., Sage, G.K., and Allendorf, F.W. 2004. Founding events influence genetic population structure of sockeye salmon (*Oncorhynchus nerka*) in Lake Clark, Alaska. Molecular Ecology 13: 277-290.

Rannala, B. and Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences USA 94: 9197-9201.

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

Richardson, L.R. and Gold, J.R. 1995a. Evolution of the *Cyprinella lutrensis* species group .III. Geographic variation in the Mitochondrial DNA of *Cyprinella lutrensis* - the influence of Pleistocene glaciation on population dispersal and divergence. Molecular Ecology 4: 163-171.

Richardson, L.R. and Gold, J.R. 1995b. Evolution of the *Cyprinella lutrensis* species complex .II. Systematics and biogeography of the Edwards plateau shiner, *Cyprinella lepida*. Copeia 1: 28-37.

Robbins, T.W. and Watts D.C. 1981. A history of shad culture with emphasis on federal hatchery systems. Pages 1-7 *in* Proceedings of 1981 American shad workshop-culture, transportation, and marking. U.S. Fish and Wildlife Service, Lamar Information Leaflet Number 82-01, Lamar, Pennsylvania.

Rosenberg, N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular Ecology Notes 4: 137-138.

Rossum, F.V. Vekemans, X.,Gratia, E., and Meerts, P. 2003. A comparative study of declining allozyme variaiton of peripheral populations of *Silene nutans L.* (Caryophyllaceae) from Westerns Europe: implications for conservation. Plant systematics and Evolution 242: 49-61.

Rottiers, D.V., Redell, L.A., Booke H.E, and Amaral S. 1992. Differences in stocks of American shad from the Columbia and Delaware rivers. Transactions of the American Fisheries Society 121: 132-136.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145: 1219-1228.

Rousset, F. 2007. Genepop'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Molecular Ecology Resources 8: 103-106.

Ruzzante, D.E. 1998. A comparison of several measures of genetic distance and population structure with microsatellite data: bias and sampling variance. Canadian Journal of Fisheries and Aquatic Sciences 55: 1-14.

Ryman, N. and Palm, S. 2006. POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Molecular Ecology Notes 6: 600-602.

Schoen, D.J., and Brown A.H.D. 1993. Conservation of allelic richness in wild crop relatives is aided by assessment of genetic markers. Proceedings of the National Academy of Sciences USA. 90: 10623-10627.

Schlötterer, C. 2002. A microsatellite-based multilocus screen for the identification of local selective sweeps. Genetics 160: 753-763.

Schlötterer, C. and Deiringer, D. 2005. A novel test statistic for the identification of local selective sweeps based on microsatellite gene diversity. *In* Selective Sweep. *Edited by* D. Nurminsky. Kluwer Academic/Plenum publishers, New York pp. 55-64.

Schmidt, R. E. 1986. Zoogeography of the Northern Appalachians. *In* Zoogeography of North American Freshwater Fishes. Hocutt, C. and Wiley, E. O. eds, John Wiley and Sons. New York. Pg. 137-159.

Seaman, A.A. 2004. Late Pleistocene history of New Brunswick, Canada. In: Ehlers, J., Gibbard, P.L. (Eds.), Quaternary Glaciations-Extent and Chronology, Part II, pp.151-167.

Seeb, J. E, Habicht, E., Olsen, J. B., Bentzen, P., Shaklee, J. B., and Seeb, L. W. 1998. Allozyme, mtDNA, and microsatellite variants describe structure of populations of pink and sockeye salmon in Alaska. North Pacific Anadromous Fish Commission Bulletin No. 1.

Shoubridge, E.A. and Leggett, W.C. 1978. Occurrence and adaptive significance of distinct reproductive strategies in local populations of American shad, p. 25-73. *In* Genetic and reproductive variation in American shad. Final Rep. Proj. AFC 10 (Connecticut) National Marine Fisheries Service U.S. 74 p.

Smith, B.R., Herbinger, C.M., and Merry, H.R. 2001a. Accurate partition of individuals into full-sib families from genetic data without parental information. Genetics 158: 1329-1338.

Smith, C.T., Nelson, R.J., Wood, C.C., and Koop, B.F. 2001b. Glacial biogeography of North American coho salmon (*Oncorhynchus kisutch*). Molecular Ecology 10: 2775-2785.

Smith, C.T., Nelson, R.J., Pollard, S., Rubidge, E., Mckay, S.J., Rodzen, J., May, B., and Koop, B. 2002. Population genetic analysis of white sturgeon (*Acipenser transmontanus*) in the Fraser River. Journal of Applied Ichthyology 18: 307-312.

Smouse, P.E. and Chevillon, C. 1998. Analytical aspects of population-specific DNA fingerprinting for individuals. Journal of Heredity 89: 143-150.

Snyder, T.P., Larsen, R.D., and Bowen, S.H. 1992. Mitochondrial DNA diversity among Lake Superior and inland lake ciscoes (*Coregonus artedi* and *C. hoyi*). Canadian Journal of Fisheries and Aquatic Sciences 49: 1902-1907.

Spidle, A.P., Kalinowski, S.T., Lubinski, B.A., Perkins, D.L., Beland, K.F., Kocik, J.F., and King, T.L. 2003. Population structure of Atlantic salmon in Maine with reference to populations from Atlantic Canada. Transactions of the American Fisheries Society 132: 196-209.

SPSS, Inc. 2000. SYSTAT version 10.0. SPSS, Inc. Chicago, Il.

St. Pierre, R.A. 2003. A case history: American shad restoration on the Susquehanna River. American Fisheries Society Symposium 35: 315-321.

StefenonV.M., Behling, H., Gailing, O.,and Finkeldey, R. 2008. Evidences of delayed size recovery in *Araucaria angustifolia* populations after post-glacial colonization of highlands in Southeastern Brazil. Annals of the Brazilian Academy of Sciences 80(3): 433-443.

Stellwag, E.J., Payne, E.S., and Rulifson, R.A. 1994. Mitochondrial DNA diversity of Roanoke River striped bass. Transactions of the American Fisheries Society 123: 321-334.

Stepien, C.A. and Faber, J.E. 1998. Population genetic structure, phylogeography and spawning philopatry in walleye (*Stizostedion vitreum*) from mitochondrial DNA control region sequences. Molecular Ecology 7: 1757-1769.

Stepien, C.A., Murphy, D.J., and Strange, R.M. 2007. Broad- to fine-scale population genetic patterning in the smallmouth bass *Micropterus dolomieu* across the Laurentian Great Lakes and beyond: an interplay of behaviour and geography. Molecular Ecology 16: 1605-1624.

Stewart, I.J., Quinn, T.P., and Bentzen, P. 2003. Evidence for fine-scale natal homing among island beach spawning sockeye salmon, *Oncorhynchus nerka*. Environmental Biology of Fishes 67: 77-85.

Storz, J.F. and Beaumont, M.A. 2002. Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. Evolution 56: 154-166.

Stott, W., Todd, T.N., and Kallemeyn, L. 2004. Genetic variability among lake whitefish from Isle Royale and the Upper Great Lakes. Annales Zoologici Fennici 41: 51-59.

Talbot, G. B. and Sykes, J. E. 1958. Atlantic coast migrations of American shad. U.S. Fisheries and Wildlife Service Fishery Bulletin 58:473-480.

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

Taylor, E.B., Pollard, S., and Louie, D. 1999. Mitochondrial DNA variation in bull trout (*Salvelinus confluentus*) from northwestern North America: implications for zoogeography and conservation. Molecular Ecology 8: 1155-1170.

Taylor, E.B. and McPhail, J.D. 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.

Tessier, N. and 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.

Thomas, C.D. 2005. Recent evolutionary effects of climate change. Pages 75-88 *in* T.E. Lovejoy and L. Hannah, editors. Climate change and biodiversity. Yale University Press, New Haven, Connecticut.

Thompson, J.N. 1998. Rapid evolution as an ecological process. Trends in Ecology & Evolution **13**: 329-332.

Tibbets, C.A. and Dowling, T.E. 1996. Effects of intrinsic and extrinsic factors on population fragmentation in three species of North American minnows (Teleostei: *Cyprinidae*). Evolution 50: 1280-1292.

Todd, T.N. and Hatcher, C.O. 1993. Genetic variability and glacial origins of yellow perch (*Perca flavescens*) in North America. Canadian Journal of Fisheries and Aquatic Sciences 50: 1828-1834.

Toline, C.A. and Baker, A.J. 1994. Genetic differentiation among populations of the northern redbelly dace (*Phoxinus eos*) in Ontario. Canadian Journal of Fisheries and Aquatic Sciences 51: 1218-1228.

Toline, C.A. and Baker, A.J. 1995. Mitochondrial DNA variation and population genetic structure of the northern redbelly dace (*Phoxinus eos*). Molecular Ecology 4: 745-753.

Turgeon, J. and Bernatchez, L. 2001. Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei: Coregoninae). Evolution 55: 2274-2286.

Vaha, J.P., Erkinaro, J., Niemela, E., and Primmer, C.R. 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. Molecular Ecology 16:2638-2654.

Väli, U., Einarsson, A., Waits, L., and Ellegren, H. 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Molecular Ecology 17: 3808-3817.

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

Vasemägi, A. 2006. The adaptive hypothesis of clinal variation revisited: Single-locus clines as a result of spatially restricted gene flow. Genetics 173: 2411-2414.

Verspoor, E., O'Sullivan, M., Arnold, A.L., Knox, D., and Amiro, P.G. 2002. Restricted matrilineal gene flow and regional differentiation among Atlantic salmon (*Salmo salar* L.) populations within the Bay of Fundy, eastern Canada. Heredity 89: 465-472.

Vilas, C., San Miguel, E., Amaro, R., and Garcia, C. 2006. Relative contribution of inbreeding depression and eroded adaptive diversity to extinction risk in small populations of shore Campion. Conservation Biology 20: 229-238.

Walburg, C. H. and Nichols, P. R. 1967. Biology and management of the American shad and status of the fisheries, Atlantic coast of the United States, 1960. U.S. Fish and Wildlife Service Special Science Report of Fisheries 550:1-105.

Waldman, J.R. and Wirgin, I.I. 1994. Origin of the present Delaware River striped bass population as shown by analysis of mitochondrial DNA. Transactions of the American Fisheries Society 123: 15-21.

Waldman, J.R., Nolan, K., Hart, J., and Wirgin, I.I. 1996. Genetic differentiation of three key anadromous fish populations of the Hudson River. Estuaries 19: 759-768.

Waldman, J.R., Grunwald, C., Stabile, J., and Wirgin, I. 2002. Impacts of life history and biogeography on the genetic stock structure of Atlantic sturgeon *Acipenser oxyrinchus oxyrinchus*, Gulf sturgeon *A. oxyrinchus desotoi*, and shortnose sturgeon *A. brevirostrum*. Journal of Applied Ichthyology 18: 509-518.

Walther, B.D., Thorrold ,S.R., and Olney, J.E. 2008. Geochemical signatures in otoliths record natal origins of American shad. Transactions of the American Fisheries Society 137: 57-69.

Waples, R.S. 1990. Temporal changes of allele frequency in Pacific salmon: implications for mixed-stock fishery analysis. Canadian Journal of Fisheries and Aquatic Sciences 47: 968-976.

Waples, R.S., and D.J. Teel. 1990. Conservation genetics of pacific salmon: 1. Temporal changes in allele frequency. Conservation Biology 4(2): 144-156.

Waples, R.S. and Gaggiotti, O. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology 15: 1419-1439.

Ward, R.D., Woodwark, M., and Skibinski, D.O.F. 1994. A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. Journal of Fish Biology 44: 213-232.

Waters, J.M., Epifanio, J.M., Gunter, T., and Brown, B.L. 2000. Homing behaviour facilitates subtle genetic differentiation among river populations of *Alosa sapidissima*: microsatellites and mtDNA. Journal of Fish Biology 56: 622-636.

Weaver, L.A., Fisher, M.T., Bosher, B.T., Claud, M.L., and Koth, L.J. 2003. Boshers Dam vertical slot fishway: a useful tool to evaluate American shad recovery efforts in the Upper James River. American Fisheries Society Symposium 35: 339-347.

Weihs, D. 1978. Tidal stream transport as an efficient method for migration. Journal for the Conservation and International Exploration of the Sea 38: 92-99.

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

White, M.M. 2000. Genetic variation in white bass. Transactions of the American Fisheries Society 129: 879-885.

Wilson, C.C., Hebert, P.D.N., Reist, J.D., and Dempson, J.B. 1996. Phylogeography and postglacial dispersal of arctic charr *Salvelinus alpinus* in North America. Molecular Ecology 5: 187-197.

Wilson, C.C. and Hebert, P.D.N. 1996. Phylogeographic origins of lake trout (*Salvelinus namaycush*) in eastern North America. Canadian Journal of Fisheries and Aquatic Science 53: 2764-2775.

Wilson, C.C. and Hebert, P.D.N. 1998. Phylogeography and postglacial dispersal of lake trout (*Salvelinus namaycush*) in North America. Canadian Journal of Fisheries and Aquatic Sciences 55: 1010-1024.

Winans, G.A., McHenry, M.L., Baker, J., Elz, A., Goodbla, A., Iwamoto, E., Kuligowski, D., Miller,K.M., Small, M.P., Spruell, P., and Van Doornik, D. 2008. Genetic inventory of anadromous pacific salmonids of the Elwha River prior to dam removal. Northwest Science 82: 128-141.

Wirgin, I., Maceda, L., Waldman, J.R., and Crittenden, R.N. 1993a. Use of Mitochondrial DNA polymorphisms to estimate the relative contributions of the Hudson River and

Chesapeake Bay striped bass stocks to the mixed fishery on the Atlantic Coast. Transactions of the American Fisheries Society 122: 669-684.

Wirgin, I.I., Ong, T.L., Maceda, L., Waldman, J.R., Moore, D., and Courtenay, S. 1993b. Mitochondrial DNA variation in striped bass (*Morone saxatilis*) from Canadian Rivers. Canadian Journal of Fisheries and Aquatic Sciences 50: 80-87.

Wirgin, I., Jessop, B., Courtenay, S., Pedersen, M., Maceda, S., and Waldman, J.R. 1995. Mixed stock analysis of striped bass in two rivers of the Bay of Fundy as revealed by mitochondrial DNA. Canadian Journal of Fisheries and Aquatic Sciences 52: 961-970.

Wirgin, I., Maceda, L., Stabile, J., and Mesing, C. 1997. An evaluation of introgression of Atlantic coast striped bass mitochondrial DNA in a Gulf of Mexico population using formalin-preserved museum collections. Molecular Ecology 6: 907-916.

Wirgin, I., Waldman, J.R., Rosko, J., Gross, R., Collins, M.R., Rogers, S.G., and Stabile, J. 2000. Genetic structure of Atlantic sturgeon populations based on mitochondrial DNA control region sequences. Transactions of the American Fisheries Society 129: 476-486.

Wirgin, I., Grunwald, C., Carlson, E., Stabile, J., Peterson, D.L., and Waldman, J. 2005. Range-wide population structure of shortnose sturgeon *Acipenser brevirostrum* based on sequence analysis of the mitochondrial DNA control region. Estuaries 28: 406-421.

Wirgin, I., Grunwald, C., Stabile, J., and Waldman, J. 2007. Genetic evidence for relict Atlantic sturgeon stocks along the mid-Atlantic coast of the USA. North American Journal of Fisheries Management 27: 1214-1229.

Wood, C.C., Riddell, B.E., Rutherford, D.T., and Withler, R.E. 1994. Biochemical genetic survey of sockeye salmon (*Oncorhynchus nerka*) in Canada. Canadian Journal of Fisheries and Aquatic Sciences 51: 114-131.

Appendix 1 List of species included in the literature survey, including marker types, genomic regions examined (mtDNA), analytical methodology used, number of populations surveyed, and the studies from which these data were obtained.

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Acipenser oxyrhynchus	Atlantic sturgeon	mtDNA	D-loop	SSCP	19	Wirgin et al. 2000, 2007; Waldman et al. 2002; Grunwald et al. 2008; Peterson et al. 2008
		Microsatellites			7	King et al. 2001
Acipenser transmontanus	White sturgeon	mtDNA	Whole genome	RFLP	2	Brown et al. 1992; Smith et al. 2002
Acipenser brevirostrum	Shortnose sturgeon	mtDNA	D-loop	SSCP	16	Waldman et al. 2002; Grunwald et al. 2002; Quattro et al. 2002; Wirgin et al. 2005
Acipenser medirostris	Green sturgeon	Microsatellites			4	Israel et al. 2004
Acipenser fluvescens	Lake sturgeon	mtDNA	D-loop	SSCP	11	DeHaan et al. 2006
		mtDNA	Whole genome	RFLP	12	Ferguson et al. 1993; Guenette et al. 1993

Species	Common name	Marker type	Genomic region	Analytical method	N	References
		Microsatellites			17	McQuown et al. 2003; DeHann et al. 2006
Agosia chrysogaster	Longfin dace	mtDNA	Whole genome	RFLP	7	Tibbets and Dowling 1996
Rhinichthys osculus	Speckled dace	mtDNA	Cytochrome b	SSCP	7	Pfrender et al. 2004
Rhinichthys cataractae	Longnose dace	mtDNA	Cytochrome b	SSCP	32	Girard and Angers 2006
Alosa sapidissima	American shad	mtDNA	Whole genome	RFLP	20	Bentzen et al. 1989; Nolan et al. 1991; Epifanio et al. 1995; Brown et al. 1996; Waters et al. 2000
		Microsatellites			33	Brown et al. 2000; Waters et al. 2000; Hasselman et al. unpubl. data
Amia calva	Bowfin	mtDNA	Whole genome	RFLP	13	Bermingham and Avise 1986
Amerius nebulosus	Brown bullhead	mtDNA	Whole genome	RFLP	12	Murdoch and Hebert 1997
Catostomus commersoni	White sucker	mtDNA	Whole genome	RFLP	11	Lafontaine and

Species	Common name	Marker type	Genomic region	Analytical method	N	References
						Dodson 1997
Coregonus autmnalis	Arctic cisco	mtDNA	Whole genome	RFLP	9	Bernatchez et al. 1991
Coregonus clupeaformis	Lake whitefish	mtDNA	Whole genome	RFLP	62	Bernatchez and Dodson 1991, 1994; Lu et al. 2001
		Microsatellites			14	Lu et al. 2001; Stott et al. 2004
Coregonus nasus	Broad whitefish	mtDNA	Whole genome	RFLP	3	Bernatchez et al. 1991
		mtDNA	Cytochrome b	SSCP	2	Patton et al. 1997
		Microsatellites			2	Patton et al. 1997
Coregonus laurettae	Bering cisco	mtDNA	Whole genome	RFLP	1	Bickham et al. 1992
Coregonus artedi	Cisco	mtDNA	Whole genome	RFLP	16	Bernatchez and Dodson 1990; Snyder et al. 1992
		mtDNA	D-loop	SSCP	27	Turgeon and Bernatchez 2001
		Microsatellites			22	Turgeon and Bernatchez 2001

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Culaea inconstans	Brook stickleback	mtDNA	Whole genome	RFLP	32	Gach 1996
Cyprinella lepida	Plateau shiner	mtDNA	Whole genome	RFLP	3	Richardson and Gold 1995b
Cyprinella lutrensis	Red shiner	mtDNA	Whole genome	RFLP	14	Richardson and Gold 1995a; Richardson and Gold 1995b
Cyprinella caerulea	Blue shiner	mtDNA	ND2	RFLP	4	George et al. 2008
Fundulus heteroclitus	Mummichog	mtDNA	Whole genome	RFLP	4	Gonzales-Villasenor and Powers 1990
		mtDNA	Whole genome	SSCP	2	Mulvey et al. 2003
		Allozyme			19	Cashon et al. 1981
		Microsatellites			38	Adams et al. 2006; Duvernell et al. 2008
Gasterosteus aculeatus	Threespine stickleback	mtDNA	Whole genome	RFLP	30	O'Reilly et al. 1993; Taylor and McPhail 1999
Lepomis punctatus	Spotted sunfish	mtDNA	Whole genome	RFLP	13	Bermingham and Avise 1986
		Allozyme			11	McElroy et al. 2003

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Lepomis punctatus	Spotted sunfish	Microsatellites			11	McElroy et al. 2003
Lepomis gibbosus	Pumpkinseed sunfish	Allozyme			4	Fox et al. 1997
Lepomis gulosus	Warmouth	mtDNA	Whole genome	RFLP	12	Bermingham and Avise 1986
Lepomis microlophus	Redear sunfish	mtDNA	Whole genome	RFLP	11	Bermingham and Avise 1986
Micropterus salmoides	Largemouth bass	mtDNA	Whole genome	RFLP	6	Nedbal and Phillip 1994
		Allozyme			89	Phillip et al. 1985
		Microsatellites			13	Lutz-Carrillo et al. 2006
Micropterus dolomieu	Smallmouth bass	Microsatellites			28	Stepien et al. 2007
Morone Americana	White perch	mtDNA	Whole genome	RFLP	7	Mulligan and Chapman 1989
		Allozyme			19	White 2000
Morone saxatilis	Striped bass	mtDNA	Whole genome	RFLP	10	Wirgin et al. 1993a,b, 1995,1997; Waldman et al. 1996; Waldman and Wirgin 1994; Stellwag et al. 1994

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Morone saxatilis	Striped bass	Allozyme			7	Bulak et al. 2004
		Microsatellites			7	Laughlin et al. 1996; Diaz et al. 1997
Oncorhynchus nerka	Sockeye salmon	mtDNA	Cytochrome b	RFLP	4	Bickham et al. 1995
		Allozyme			83	Wood et al. 1994
		Microsatellites			34	Seeb et al. 1998; Allendorf and Seeb 2000; Beacham et al. 2002; Stewart et al. 2003; Ramstad et al. 2003, 2004; Olsen et al. 2004; Habicht et al. 2007; Lin et al. 2008; Winans et al. 2008
Oncorhynchus mykiss	Rainbow trout	mtDNA	D-loop	SSCP	12	Bagley and Gall 1998
		mtDNA	Whole genome	SSCP/ RFLP	11	McCusker et al. 2000
		Microsatellites			2	Deiner et al. 2007; Pearse et al. 2007
Oncorhynchus kisutch	Coho salmon	mtDNA	D-loop	SSCP	17	Smith et al. 2001b

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Oncorhynchus kisutch	Coho salmon	Microsatellites			24	Johnson and Banks 2008
Oncorhynchus keta	Chum salmon	Allozyme			8	Phelps et al. 1994
Oncorhynchus tshawytscha	Chinook salmon	Microsatellites			15	Banks et al. 2000; Beacham et al. 2003
Osmerus mordax	Rainbow smelt	mtDNA	Whole genome	RFLP	42	Taylor and Bentzen 1993; Bernatchez and Martin 1996; Baby et al. 1991 as in Bernatchez 1997
		Microsatellites			21	Bradbury et al. 2006
Phoxinus eos	Northern redbelly dace	mtDNA	ND5,6	RFLP	2	Toline and Baker 1995
		Allozymes			3	Toline and Baker 1994
Poecilipsis occidentalis	Sonoran topminnow	mtDNA	Whole genome	RFLP	9	Quattro et al. 1996
		Microsatellites			4	Parker et al. 1999
Polyodon spathula	Mississippi paddlefish	mtDNA	Whole genome	RFLP	8	Epifanio et al. 1996
		Microsatellites			12	Heist and Mustapha 2008

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Salmo salar	Atlantic salmon	mtDNA	Whole genome	RFLP	29	King et al. 2000
		mtDNA	ND1	SSCP	11	Verspoor et al. 2002
		Microsatellites			43	McConnell et al. 1997; Tessier and Bernatchez 1999; King et al. 2001; Spidle et al. 2003
Salvelinus alpines	Arctic charr	mtDNA	Whole genome	RFLP	53	Wilson et al. 1996
		Microsatellites			23	Bernatchez et al. 1998, 2002; Lundrigan et al. 2005
Salvelins namaycush	Lake trout	mtDNA	Whole genome	RFLP	93	Wilson and Hebert 1996, 1998
Salvelinus fontinalis	Brook charr	mtDNA	Whole genome	RFLP	97	Danzmann et al. 1998
		Microsatellites			51	Castric et al. 2001; Fraser and Bernatchez 2005
Sander vitreus	Walleye	mtDNA	Whole genome	RFLP	13	Stepien and Faber 1998; McParland 1999

Species	Common name	Marker type	Genomic region	Analytical method	N	References
Sander vitreus	Walleye	Allozyme			7	McParland 1999
		Microsatellites			46	Cena et al. 2006
Xyrauchon texanus	Razorback sucker	mtDNA	Whole genome	RFLP	5	Dowling et al. 1996
Salvelinus confluentus	Bull trout	mtDNA	Whole genome	RFLP	37	Taylor et al. 1999
Coregonus hoyi	Bloater	Microsatellites			12	Fave and Turgeon 2008
Perca flavescens	Yellow perch	Allozyme			13	Todd and Hatcher 1993

Appendix 2 Pearson correlations for indices of intraspecific genetic variation with latitude among 42 species of Nearctic fishes

	h		π		H_O		R
Species	RFLP	SSCP	RFLP	SSCP	Allozyme	Microsatellite	Microsatellite
Acipenser oxyrhynchus		-0.857[a]		0.981[a]		-0.720	-0.149
Acipenser brevirostrum		0.070[a]		0.283[a]			
Acipenser medirostris						-0.103	
Acipenser fluvescens	-0.223	0.169[a]				-0.444	-0.428
Agosia chrysogaster	-0.690		-0.754				
Rhinichthys osculus				-0.447[b]			
Rhinichthys cataractae		-0.136[b]		-0.009[b]			
Alosa sapidissima	-0.426					-0.648	-0.873
Amerius nebulosus	0.134		0.074				
Catostomus commersoni	0.247						
Coregonus autmnalis	0.561						
Coregonus clupeaformis-dwarf (Acadian refugium)	0.169		-0.083			-0.055	
Coregonus clupeaformis-normal (Acadian refugium)	0.290		0.326			0.859	
Coregonus clupeaformis- ecotypes pooled (Acadian refugium)	0.286		0.059				
Coregonus clupeaformis-normal (Atlantic refugium)	0.360		-0.482				
Coregonus clupeaformis-normal (Beringian refugium)			0.787				
Coregonus clupeaformis-normal (Mississippian refugium)			0.496			-0.057	
Coregonus clupeaformis-normal (refugia pooled)	0.274		0.536			0.249	
Coregonus artedi	0.307	-0.193		-0.262[a]		0.554	

	h		π		H_O		R
Species	RFLP	SSCP	RFLP	SSCP	Allozyme	Microsatellite	Microsatellite
Cyprinella lepida	-0.834						
Cyprinella lutrensis	-0.572	0.563					
Cyprinella caerulea		0.321 [c]					
Fundulus heteroclitus	-0.667				-0.614	-0.692	-0.828
Gasterosteus aculeatus	0.211		0.402				
Lepomis punctatus					0.096		
Lepomis gibbosus					-0.711		
Micropterus salmoides	-0.336				-0.404	0.611	
Micropterus dolomieu						-0.569	
Morone americana	-0.744				0.120		
Morone saxatilis	-0.746				0.691	-0.747	
Oncorhynchus nerka	-0.630 [b]				0.148	-0.380	0.115
Oncorhynchus mykiss	-0.367	-0.012	-0.037				
Oncorhynchus kisutch		-0.797		-0.811		0.026	
Oncorhynchus keta					-0.439		
Oncorhynchus tshawytscha						0.898	
Osmerus mordax-glacial lineage 'A'	0.311		-0.269			-0.691	
Osmerus mordax-glacial lineage 'B'	0.215						
Phoxinus eos					0.492		
Poecilipsis occidentalis						-0.851	
Polyodon spathula						0.330	
Salmo salar	0.092	0.434 [c]				-0.143	
Salvelinus alpinus	-0.562					-0.026	-0.539
Salvelinus namaycush - Atlantic refugium	0.176		0.227				

	h		π		H_O		R
Species	RFLP	SSCP	RFLP	SSCP	Allozyme	Microsatellite	Microsatellite
Salvelinus namaycush - Mississippian refugium	0.078		0.072				
Salvelinus namaycush - Beringian refugium	0.719		-0.777				
Salvelinus namaycush - Nahanni refugium	-0.017		0.098				
Salvelinus fontinalis	-0.277					0.154	0.919
Sander vitreus	-0.442				-0.623	-0.189	
Xyrauchon texanus	-0.600						
Salvelinus confluentus	0.737		0.655				
Coregonus hoyi						0.765	-0.424
Perca flavescens					-0.473		

a - D-loop
b - Cytochrome b
c - ND genes

Appendix 3 (back pocket) Genetic diversity statistics for American shad collections in Atlantic Canada from 2003-2006. Statistics per locus across collections include number of alleles (A), and range in allele size. Statistics per collections across loci include allelic richness (R; standardized n=32), expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}). Statistics per collection and per locus include sample size (N), number of alleles (N_a), allelic richness (R), F_{IS}, H_E, and observed heterozygosity (H_O).

Appendix 3

			STL		MR			RP	MUS	LH	AR	ANN	GR		KR	SHU			RH	SJR		
Locus	A	Year	'04	'05	'04	'05	'06	'04-'06	'04-'06	'04-'06	'04-'06	'04	'04	'06	'05	'04	'05	'06	'04-'05	'03	'04	'05	'06
Asa2	24	N	81	204	122	173	32	85	43	67	17	99	80	35	50	272	246	227	45	76	25	121	214
Range	73-160	N_a	18	20	16	18	15	18	10	13	8	14	14	9	13	17	19	17	10	14	10	16	16
		R	13.365	11.504	13.181	13.168	NA	11.977	8.717	11.153	NA	11.769	10.840	8.907	11.090	12.042	13.268	12.035	7.061	12.033	NA	11.900	11.263
		F_{IS}	0.021	-0.038	0.009	0.004	0.125	0.032	-0.030	0.014	-0.096	0.031	0.070	0.046	-0.100	0.035	0.022	0.047	-0.009	0.026	0.016	0.043	-0.006
		H_E	0.820	0.793	0.852	0.847	0.856	0.839	0.768	0.787	0.807	0.865	0.739	0.689	0.801	0.850	0.852	0.846	0.661	0.756	0.731	0.760	0.734
		H_O	0.803	0.824	0.844	0.844	0.750	0.812	0.791	0.776	0.882	0.838	0.688	0.657	0.880	0.820	0.833	0.806	0.667	0.737	0.720	0.727	0.738
Asa4	14	N	96	209	146	187	33	82	42	67	17	105	82	35	50	280	246	228	45	89	23	117	210
Range	125-164	N_a	12	9	11	9	8	8	8	8	8	6	9	7	8	10	9	9	7	10	6	10	11
		R	9.118	7.002	8.901	7.307	NA	7.614	7.705	7.639	NA	5.998	7.885	6.994	7.511	8.173	8.496	8.125	7.061	7.833	NA	7.855	7.868
		F_{IS}	0.032	-0.032	0.047	-0.077	-0.132	-0.113	0.003	0.023	0.046	-0.012	-0.009	0.150	-0.120	0.018	-0.001	0.007	0.012	0.170	0.093	0.002	0.020
		H_E	0.818	0.709	0.747	0.685	0.644	0.724	0.788	0.794	0.800	0.810	0.822	0.805	0.805	0.796	0.796	0.777	0.787	0.676	0.574	0.677	0.651
		H_O	0.792	0.732	0.712	0.738	0.727	0.805	0.786	0.776	0.765	0.819	0.829	0.686	0.900	0.782	0.797	0.772	0.778	0.562	0.522	0.675	0.638
Asa8	12	N	71	210	155	188	32	80	40	65	17	99	82	34	48	276	246	229	44	73	23	119	211
Range	108-156	N_a	5	7	8	7	6	7	7	7	7	7	7	8	7	7	8	8	7	9	7	8	9
		R	4.953	4.456	5.739	5.430	NA	6.229	6.761	6.142	NA	6.766	6.918	7.936	6.333	6.112	6.560	6.474	7.061	7.384	NA	6.842	6.507
		F_{IS}	0.170	0.098	-0.051	0.048	0.050	0.023	-0.137	0.095	0.054	0.032	0.116	-0.160	-0.064	0.023	0.087	0.018	-0.053	0.111	0.050	-0.087	-0.012
		H_E	0.746	0.681	0.663	0.643	0.690	0.704	0.749	0.714	0.745	0.824	0.814	0.813	0.803	0.801	0.792	0.796	0.777	0.739	0.777	0.750	0.735
		H_O	0.620	0.614	0.697	0.612	0.656	0.688	0.850	0.646	0.706	0.798	0.720	0.941	0.854	0.783	0.724	0.782	0.818	0.658	0.739	0.815	0.744
Asa16	8	N	82	205	153	186	31	79	42	68	16	104	82	34	50	270	244	227	44	64	24	121	213
Range	103-154	N_a	3	4	5	5	3	7	5	7	6	6	4	4	4	5	4	5	3	4	5	6	5
		R	2.630	2.312	3.683	3.291	NA	6.015	4.999	5.936	NA	4.610	3.919	4.000	3.280	4.094	3.665	3.755	7.061	4.000	NA	4.764	4.317
		F_{IS}	-0.018	0.008	-0.006	-0.040	-0.075	0.047	-0.079	0.094	-0.193	-0.037	-0.023	-0.173	0.057	-0.024	0.061	-0.059	-0.036	-0.040	0.143	0.091	-0.002
		H_E	0.335	0.334	0.227	0.212	0.180	0.452	0.707	0.730	0.685	0.501	0.429	0.478	0.318	0.452	0.476	0.445	0.461	0.586	0.582	0.473	0.572

			STL			MR		RP	MUS	LH	AR	ANN	GR		KR		SHU		RH		SJR		
Locus	A	Year	‘04	‘05	‘04	‘05	‘06	‘04-‘06	‘04-‘06	‘04-‘06	‘04-‘06	‘04	‘04	‘06	‘05	‘04	‘05	‘06	‘04-‘05	‘03	‘04	‘05	‘06
		H_O	0.342	0.332	0.229	0.220	0.194	0.430	0.762	0.662	0.813	0.519	0.439	0.559	0.300	0.463	0.447	0.471	0.477	0.609	0.500	0.430	0.573
Aa14	39	N	93	214	165	191	33	86	43	67	17	102	81	34	47	274	240	219	44	83	23	110	207
Range	114-194	N_a	11	17	20	27	11	10	9	9	8	21	20	8	14	23	24	26	7	13	11	21	27
		R	7.210	8.087	10.099	13.044	NA	15.702	21.996	16.590	NA	13.191	12.171	7.995	11.784	13.911	14.613	15.285	7.061	11.036	NA	13.642	14.124
		F_{IS}	0.161	0.080	0.070	0.149	-0.083	0.109	0.117	0.142	0.043	0.117	-0.014	-0.056	0.071	0.041	-0.011	-0.036	-0.006	-0.003	0.236	0.055	0.003
		H_E	0.486	0.569	0.612	0.738	0.756	0.796	0.868	0.817	0.859	0.821	0.780	0.809	0.709	0.712	0.754	0.750	0.565	0.709	0.679	0.711	0.775
		H_O	0.409	0.523	0.570	0.628	0.818	0.709	0.767	0.702	0.824	0.726	0.790	0.853	0.660	0.683	0.763	0.776	0.568	0.711	0.522	0.673	0.773
Af6	26	N	71	204	93	176	29	75	42	64	17	90	75	35	48	258	182	196	44	76	20	106	169
Range	143-195	N_a	9	10	14	14	9	19	13	10	9	12	17	8	12	23	19	21	12	15	10	20	20
		R	7.287	6.876	10.820	8.446	NA	14.200	11.924	9.184	NA	9.011	15.072	7.902	10.586	14.059	12.988	12.585	7.061	12.257	NA	14.453	11.025
		F_{IS}	0.321	-0.054	0.142	-0.017	0.008	0.063	0.083	-0.033	0.018	0.071	0.059	-0.010	0.028	0.056	0.008	0.048	0.137	0.082	0.229	0.019	-0.012
		H_E	0.683	0.563	0.801	0.654	0.695	0.797	0.804	0.772	0.718	0.753	0.878	0.793	0.835	0.862	0.864	0.847	0.815	0.889	0.838	0.875	0.853
		H_O	0.465	0.593	0.688	0.665	0.690	0.747	0.738	0.797	0.706	0.700	0.827	0.800	0.813	0.814	0.857	0.806	0.705	0.816	0.650	0.859	0.864
Af13	13	N	102	209	142	185	33	81	42	64	17	97	81	34	50	273	233	224	42	85	14	106	199
Range	160-186	N_a	7	6	8	8	8	9	7	7	6	8	9	7	8	10	8	9	6	11	7	8	7
		R	6.754	5.409	7.153	6.675	NA	7.844	6.749	6.245	NA	7.694	8.175	6.997	7.509	8.206	6.722	6.557	7.061	9.012	NA	6.477	6.535
		F_{IS}	0.157	0.228	0.038	0.155	0.235	0.096	0.203	0.128	0.141	0.151	0.017	-0.032	0.381	0.179	0.131	0.140	-0.058	0.039	0.403	-0.036	0.073
		H_E	0.778	0.632	0.812	0.773	0.790	0.791	0.775	0.770	0.818	0.837	0.854	0.798	0.772	0.838	0.790	0.779	0.721	0.722	0.825	0.738	0.672
		H_O	0.657	0.488	0.782	0.654	0.606	0.716	0.619	0.672	0.706	0.711	0.840	0.824	0.480	0.689	0.687	0.670	0.762	0.694	0.500	0.764	0.623
Af20	4	N	105	195	164	192	33	83	43	67	17	104	79	35	50	274	242	228	43	89	23	113	206
Range	169-175	N_a	3	3	3	3	3	4	3	4	3	3	3	2	3	3	3	3	2	4	3	4	3
		R	2.304	2.735	3.000	3.000	NA	3.909	3.000	3.477	NA	2.985	2.792	2.000	2.873	2.583	2.247	2.841	7.061	3.591	NA	3.634	3.000
		F_{IS}	-0.107	0.062	-0.038	-0.001	-0.230	0.171	0.271	0.058	-0.067	0.135	0.309	-0.193	0.064	-0.008	-0.073	-0.194	-0.200	-0.144	0.154	-0.041	-0.149
		H_E	0.189	0.131	0.523	0.520	0.470	0.508	0.541	0.491	0.221	0.333	0.402	0.288	0.427	0.391	0.374	0.426	0.291	0.600	0.614	0.587	0.545

			STL		MR			RP	MUS	LH	AR	ANN	GR		KR	SHU			RH	SJR		
Locus	A	Year	‘04	‘05	‘04	‘05	‘06	‘04-‘06	‘04-‘06	‘04-‘06	‘04-‘06	‘04	‘04	‘06	‘05	‘04	‘05	‘06	‘04-‘05	‘03	‘04	‘05	‘06
		H_O	0.210	0.123	0.543	0.521	0.576	0.422	0.395	0.463	0.235	0.289	0.279	0.343	0.400	0.394	0.401	0.509	0.349	0.685	0.522	0.611	0.626
Aps2A	16	N	103	212	165	187	32	83	42	66	17	93	79	32	40	275	226	217	41	86	17	104	204
Range	62-122	N_a	9	13	13	10	9	11	11	12	10	10	11	9	8	14	10	13	9	11	7	13	12
		R	8.027	7.836	9.037	7.429	NA	8.259	10.636	10.728	NA	8.455	9.235	9.000	7.800	11.067	8.864	9.671	7.061	9.597	NA	9.960	8.962
		F_{IS}	-0.051	-0.025	0.020	0.013	-0.007	-0.126	0.056	-0.027	0.041	0.013	-0.050	-0.070	0.026	-0.051	0.029	0.018	0.263	0.161	-0.037	0.114	0.098
		H_E	0.786	0.787	0.804	0.785	0.838	0.792	0.857	0.841	0.857	0.806	0.856	0.848	0.847	0.869	0.848	0.849	0.825	0.858	0.852	0.836	0.837
		H_O	0.825	0.807	0.788	0.775	0.844	0.892	0.810	0.864	0.824	0.796	0.899	0.906	0.825	0.913	0.823	0.834	0.610	0.721	0.882	0.740	0.755
AsaD042	17	N	71	212	123	184	33	83	40	66	17	89	81	35	50	263	244	225	45	81	21	98	200
Range	150-214	N_a	10	12	10	14	10	9	8	10	6	6	11	7	9	12	10	12	9	11	9	10	13
		R	8.531	8.791	8.519	9.883	NA	7.924	7.561	7.850	NA	5.331	8.643	6.908	8.613	9.247	8.629	9.073	7.061	9.777	NA	9.312	9.763
		F_{IS}	-0.072	-0.036	-0.061	-0.010	0.177	-0.009	-0.134	0.101	0.198	-0.085	-0.056	-0.100	-0.088	0.003	-0.083	-0.029	-0.030	-0.037	0.064	-0.065	-0.033
		H_E	0.723	0.728	0.759	0.802	0.881	0.824	0.751	0.673	0.656	0.704	0.795	0.781	0.791	0.816	0.799	0.808	0.755	0.809	0.813	0.786	0.789
		H_O	0.775	0.755	0.805	0.810	0.727	0.831	0.850	0.606	0.529	0.764	0.840	0.857	0.860	0.814	0.865	0.831	0.778	0.840	0.762	0.837	0.815
AsaB020	14	N	87	206	143	185	33	85	43	66	17	91	77	35	50	268	237	227	45	70	22	105	195
Range	115-154	N_a	10	11	8	10	10	5	6	6	6	7	7	7	7	8	9	9	6	9	7	8	10
		R	7.690	7.322	7.499	7.913	NA	6.759	9.540	7.782	NA	6.945	6.993	6.994	6.994	6.898	6.638	6.924	7.061	8.233	NA	7.277	7.960
		F_{IS}	-0.110	0.027	-0.046	-0.005	0.065	0.046	-0.095	0.129	0.142	0.092	-0.075	-0.081	0.106	-0.036	-0.069	0.006	-0.062	0.034	-0.019	0.005	-0.013
		H_E	0.798	0.764	0.842	0.828	0.842	0.826	0.829	0.816	0.888	0.750	0.834	0.847	0.827	0.800	0.813	0.798	0.775	0.799	0.803	0.785	0.790
		H_O	0.885	0.743	0.881	0.832	0.788	0.788	0.907	0.712	0.765	0.681	0.896	0.914	0.740	0.828	0.869	0.793	0.822	0.771	0.818	0.781	0.800
AsaD429	12	N	102	193	162	167	33	84	42	64	17	94	78	35	40	261	225	224	45	87	24	81	197
Range	140-196	N_a	9	8	9	9	9	9	9	7	7	6	8	7	7	8	8	8	7	9	8	9	8
		R	8.018	7.423	8.869	8.795	NA	8.001	8.705	6.848	NA	5.975	7.016	6.829	6.562	6.419	6.122	6.504	7.061	7.540	NA	7.565	7.526
		F_{IS}	0.043	-0.009	0.019	-0.043	-0.026	-0.006	-0.001	-0.149	-0.201	-0.011	0.046	-0.062	0.111	-0.075	0.096	0.006	0.007	0.059	0.022	-0.033	-0.054
		H_E	0.788	0.786	0.837	0.827	0.857	0.817	0.809	0.749	0.690	0.695	0.766	0.781	0.730	0.745	0.752	0.768	0.739	0.806	0.767	0.801	0.814

			STL		MR			RP	MUS	LH	AR	ANN	GR		KR	SHU			RH	SJR			
Locus	A	Year	‘04	‘05	‘04	‘05	‘06	‘04-‘06	‘04-‘06	‘04-‘06	‘04-‘06	‘04	‘04	‘06	‘05	‘04	‘05	‘06	‘04-‘05	‘03	‘04	‘05	‘06
		H_O	0.755	0.793	0.821	0.862	0.879	0.821	0.810	0.859	0.824	0.702	0.731	0.829	0.650	0.801	0.680	0.763	0.733	0.759	0.750	0.827	0.858
AsaD029	18	N	77	198	158	168	27	82	43	68	17	93	81	35	50	270	245	229	45	85	26	118	202
Range	156-260	N_a	10	10	11	11	9	8	8	9	7	8	10	7	8	10	11	9	8	11	8	12	12
		R	7.790	6.529	7.985	8.377	NA	7.309	7.422	7.922	NA	6.485	8.289	6.829	7.700	7.857	7.628	7.134	7.061	8.173	NA	8.229	8.188
		F_{IS}	0.020	0.011	0.022	-0.026	-0.074	0.026	-0.034	-0.116	0.068	0.008	-0.002	-0.096	-0.010	-0.030	-0.005	-0.082	-0.015	0.062	-0.121	-0.061	-0.047
		H_E	0.795	0.786	0.686	0.702	0.795	0.676	0.607	0.659	0.756	0.759	0.776	0.757	0.733	0.748	0.756	0.726	0.745	0.652	0.791	0.719	0.700
		H_O	0.779	0.778	0.671	0.720	0.852	0.659	0.628	0.735	0.706	0.753	0.778	0.829	0.740	0.770	0.759	0.786	0.756	0.612	0.885	0.763	0.733
		R	7.21	6.64	8.04	7.90	NA	8.60	8.90	8.27	NA	7.32	8.30	6.87	7.59	8.51	8.19	8.19	7.06	8.50	NA	8.61	8.23
All loci		H_E	0.673	0.636	0.705	0.694	0.715	0.734	0.758	0.739	0.731	0.728	0.750	0.730	0.723	0.745	0.744	0.740	0.686	0.738	0.742	0.731	0.728
		F_{IS}	0.050	0.019	0.015	0.015	0.020	0.024	0.014	0.036	0.024	0.038	0.020	-0.055	0.032	0.013	0.017	0.002	0.011	0.045	0.093	0.000	-0.008

Appendix 4 (back pocket) Genetic differentiation among collections of American shad within the Canadian portion of the species' range. Matrix of pairwise F_{ST} estimates (θ; Weir and Cockerham 1984) below diagonal, and standardized F'_{ST} estimates (Hedrick 2005) above diagonal (bold values indicates non-significant differentiation; $p>0.05$).

Appendix 4

Collection	STL04	STL05	MR04	MR05	MR06	RP	MUS	LH	AR	ANN	GR04	GR06	KR	SHU04	SHU05	SHU06	RH	SJR03	SJR04	SJR05	SJR06
STL04		0.0280	0.1145	0.1346	0.1413	0.1694	0.2822	0.2791	0.3198	0.2511	0.2204	0.2234	0.1851	0.1913	0.1860	0.1998	0.1973	0.2613	0.2662	0.2055	0.2577
STL05	0.0097		0.1466	0.1304	0.1469	0.1937	0.3224	0.3073	0.3427	0.2853	0.2727	0.2729	0.2296	0.2352	0.2297	0.2452	0.2451	0.3190	0.3132	0.2493	0.2935
MR04	0.0355	0.0486		0.0231	0.0400	0.0531	0.1781	0.1861	0.2287	0.2268	0.1690	0.2007	0.1195	0.1404	0.1461	0.1454	0.2053	0.1477	0.1549	0.1124	0.1552
MR05	0.0425	0.0438	0.0070		0.0114	0.0515	0.1985	0.2076	0.2523	0.2313	0.1964	0.2226	0.1432	0.1663	0.1653	0.1655	0.2199	0.1793	0.1720	0.1331	0.1735
MR06	0.0437	0.0491	0.0116	0.0034		0.0649	0.2153	0.2354	0.2619	0.2312	0.1755	0.1810	0.1305	0.1632	0.1627	0.1539	0.2100	0.1882	0.1513	0.1126	0.1533
RP	0.0503	0.0624	0.0150	0.0149	0.0178		0.0901	0.1188	0.1767	0.1722	0.1408	0.1656	0.0799	0.1070	0.1003	0.1027	0.1899	0.1091	0.1068	0.0752	0.0916
MUS	0.0817	0.1020	0.0486	0.0556	0.0566	0.0230		0.0079	0.0973	0.1789	0.1707	0.1869	0.1041	0.1373	0.1324	0.1284	0.1697	0.1478	0.1620	0.1488	0.1510
LH	0.0824	0.0987	0.0521	0.0595	0.0639	0.0312	**0.002**		0.0554	0.1492	0.1643	0.1944	0.1135	0.1339	0.1343	0.1395	0.1759	0.1827	0.2101	0.1974	0.2027
AR	0.0973	0.1134	0.0652	0.0738	0.0727	0.0472	0.0247	0.0146		0.1403	0.1712	0.1979	0.1416	0.1300	0.1357	0.1471	0.2406	0.2572	0.2636	0.2458	0.2545
ANN	0.0751	0.0923	0.0645	0.0673	0.0641	0.0463	0.0463	0.0398	0.0380		0.1273	0.1094	0.1002	0.1209	0.1228	0.1307	0.1932	0.2144	0.2240	0.1827	0.2054
GR04	0.0637	0.0860	0.0464	0.0553	0.0465	0.0363	0.0421	0.0419	0.0441	0.0333		0.0175	0.0698	0.0792	0.0900	0.0964	0.1087	0.1909	0.1427	0.1700	0.1932
GR06	0.0676	0.0897	0.0572	0.0650	0.0503	0.0444	0.0479	0.0515	0.0535	0.0297	**0.0045**		0.0794	0.1169	0.1131	0.1128	0.1045	0.2135	0.1771	0.1660	0.1810
KR	0.0563	0.0757	0.0343	0.0421	0.0366	0.0216	0.0270	0.0304	0.0387	0.0275	0.0183	0.0217		0.0201	0.0172	0.0097	0.0588	0.1699	0.1569	0.0942	0.1302
SHU04	0.0547	0.0723	0.0384	0.0465	0.0435	0.0278	0.0343	0.0345	0.0338	0.0318	0.0201	0.0306	0.0053		0.0045	0.0053	0.0936	0.1609	0.1389	0.1204	0.1662
SHU05	0.0534	0.0711	0.0401	0.0464	0.0436	0.0261	0.0332	0.0347	0.0354	0.0324	0.0228	0.0296	0.0046	0.0012		0.0000	0.0949	0.1735	0.1533	0.1106	0.1505
SHU06	0.0579	0.0765	0.0402	0.0468	0.0416	0.0270	0.0325	0.0363	0.0388	0.0347	0.0247	0.0298	0.0026	0.0014	**0.000**		0.0889	0.1663	0.1395	0.0969	0.1382
RH	0.0634	0.0846	0.0622	0.0680	0.0630	0.0546	0.0472	0.0502	0.0709	0.0561	0.0304	0.0306	0.0173	0.0261	0.0265	0.0251		0.2353	0.2253	0.1966	0.2423
SJR03	0.0771	0.1022	0.0414	0.0515	0.0512	0.0288	0.0373	0.0476	0.0679	0.0572	0.0489	0.0567	0.0457	0.0415	0.0449	0.0434	0.0671		0.0205	0.0698	0.0944
SJR04	0.0795	0.1021	0.0434	0.0495	0.0411	0.0280	0.0403	0.0544	0.0691	0.0596	0.0361	0.0469	0.0420	0.0356	0.0394	0.0362	0.0650	**0.0053**		0.0441	0.0586
SJR05	0.0611	0.0801	0.0317	0.0385	0.0310	0.0201	0.0383	0.0524	0.0662	0.0495	0.0442	0.0448	0.0257	0.0315	0.0290	0.0256	0.0568	0.0186	**0.0117**		0.0085
SJR06	0.0764	0.0934	0.0439	0.0502	0.0425	0.0247	0.0392	0.0541	0.0690	0.0559	0.0506	0.0491	0.0357	0.0437	0.0397	0.0368	0.0700	0.0252	**0.0156**	0.0023	

Appendix 5 (back pocket) Genetic diversity statistics for American shad spawning runs from across the species' range collected from 2003-2006. Statistics per locus across spawning runs include number of alleles (A), and range in allele size. Statistics per spawning run across loci include allelic richness (R; standardized n=32), expected heterozygosity (H_E), and inbreeding coefficient (F_{IS}). Statistics per spawning run and per locus include sample size (N), number of alleles (N_a), allelic richness (R), F_{IS}, H_E, and observed heterozygosity (H_O).

Appendix 5

Locus	A		STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD	DEL	NAN	SUS	PAT	POT	RAP	YOR
Asa2	26	N	285	328	85	43	67	17	99	116	50	745	45	436	80	93	89	25	12	52	52	100	58	94
Range	73-160	N_a	21	18	18	10	13	8	14	14	13	19	10	17	15	17	18	12	12	15	15	18	16	20
		R	12.14	13.48	12.12	8.84	11.25	NA	11.85	10.30	11.23	12.55	9.12	11.77	12.75	14.08	15.09	NA	NA	13.69	13.89	14.38	13.79	16.13
		F_{IS}	-0.022	0.016	0.032	-0.030	0.014	-0.096	0.031	0.062	-0.100	0.035	-0.009	0.013	0.013	0.030	0.027	0.124	0.016	-0.026	-0.011	0.025	0.117	0.013
		H_E	0.800	0.849	0.839	0.768	0.787	0.807	0.865	0.726	0.801	0.850	0.661	0.744	0.861	0.887	0.889	0.911	0.931	0.881	0.875	0.882	0.878	0.905
		H_O	0.818	0.835	0.812	0.791	0.776	0.882	0.838	0.681	0.880	0.820	0.667	0.734	0.850	0.860	0.865	0.800	0.917	0.904	0.885	0.860	0.776	0.894
Asa4	25	N	305	368	82	42	67	17	105	118	50	754	45	439	78	96	93	25	12	54	61	115	79	97
Range	119-191	N_a	12	11	8	8	8	8	6	9	8	10	7	13	9	11	12	9	8	8	11	15	10	17
		R	8.33	8.45	7.65	7.74	7.67	NA	6.00	7.63	7.55	8.31	6.73	8.13	7.93	8.83	9.31	NA	NA	7.79	9.02	9.84	8.05	10.38
		F_{IS}	-0.001	-0.028	-0.113	0.003	0.023	0.046	-0.012	0.032	-0.120	0.008	0.012	0.050	0.001	0.010	-0.026	-0.167	-0.106	-0.005	0.009	-0.021	-0.086	-0.097
		H_E	0.750	0.708	0.724	0.788	0.794	0.800	0.810	0.814	0.805	0.790	0.787	0.659	0.680	0.747	0.734	0.722	0.757	0.719	0.744	0.715	0.711	0.734
		H_O	0.751	0.728	0.805	0.786	0.776	0.765	0.819	0.788	0.900	0.784	0.778	0.626	0.680	0.740	0.753	0.840	0.833	0.722	0.738	0.730	0.772	0.804
Asa8	17	N	281	376	80	40	65	17	99	117	48	751	44	426	83	95	92	24	10	53	59	109	81	89
Range	108-172	N_a	7	8	7	7	7	7	7	8	7	9	7	11	12	13	14	8	5	9	10	12	13	10
		R	4.76	5.76	6.26	6.80	6.19	NA	6.78	7.19	6.38	6.39	6.99	6.84	9.49	9.92	10.07	NA	NA	8.24	8.97	9.95	9.14	8.32
		F_{IS}	0.120	0.011	0.023	-0.137	0.095	0.054	0.032	0.034	-0.064	0.042	-0.053	-0.005	-0.080	0.065	-0.027	0.004	0.194	0.053	0.010	0.035	0.082	-0.012
		H_E	0.700	0.659	0.704	0.749	0.714	0.745	0.824	0.814	0.803	0.796	0.777	0.745	0.804	0.822	0.794	0.753	0.737	0.836	0.771	0.799	0.780	0.755
		H_O	0.616	0.652	0.688	0.850	0.646	0.706	0.798	0.786	0.854	0.763	0.818	0.749	0.868	0.768	0.815	0.750	0.600	0.793	0.763	0.771	0.716	0.764
Asa16	13	N	287	372	79	42	68	16	104	117	50	741	44	422	83	96	93	24	11	54	53	115	83	96
Range	103-154	N_a	4	6	7	5	7	6	6	4	4	5	3	6	6	5	6	6	4	6	6	7	6	8
		R	2.42	3.52	6.07	5.00	5.97	NA	4.63	3.95	3.32	3.89	2.94	4.48	5.35	4.72	5.09	NA	NA	5.22	5.97	5.96	5.78	6.49
		F_{IS}	-0.001	-0.028	0.047	-0.079	0.094	-0.193	-0.037	-0.071	0.057	-0.006	-0.036	0.027	-0.127	-0.069	0.016	-0.169	0.320	0.071	-0.048	0.052	0.029	-0.059
		H_E	0.334	0.214	0.452	0.707	0.730	0.685	0.501	0.439	0.318	0.457	0.461	0.548	0.631	0.624	0.633	0.787	0.658	0.598	0.666	0.624	0.595	0.709
		H_O	0.335	0.220	0.430	0.762	0.662	0.813	0.519	0.470	0.300	0.460	0.477	0.533	0.711	0.667	0.624	0.917	0.455	0.556	0.698	0.591	0.578	0.750
Aa14	43	N	307	391	86	43	67	17	102	116	47	733	44	423	87	94	97	25	10	50	52	97	73	92
Range	114-202	N_a	20	31	23	25	22	11	21	22	14	31	11	31	22	27	30	17	7	16	13	23	22	23
		R	8.04	12.10	15.91	22.33	16.82	NA	13.37	11.93	11.95	14.86	10.41	13.86	14.10	18.88	20.01	NA	NA	13.99	10.47	16.81	16.45	16.54
		F_{IS}	0.104	0.102	0.109	0.117	0.142	0.043	0.117	-0.029	0.071	0.000	-0.006	0.029	0.016	-0.003	-0.022	-0.137	0.060	0.126	-0.100	0.048	-0.016	-0.071
		H_E	0.545	0.692	0.796	0.868	0.817	0.859	0.821	0.788	0.709	0.737	0.565	0.743	0.713	0.827	0.776	0.705	0.742	0.777	0.525	0.726	0.741	0.741
		H_O	0.489	0.622	0.709	0.767	0.702	0.824	0.726	0.810	0.660	0.737	0.568	0.721	0.701	0.830	0.794	0.800	0.700	0.680	0.577	0.691	0.753	0.794
Af6	33	N	275	300	75	42	64	17	90	111	48	636	44	371	72	93	88	24	11	50	56	97	67	89
Range	141-207	N_a	13	19	19	13	10	9	12	17	12	24	12	24	18	23	20	16	11	18	20	25	26	24

Locus	A		STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD	DEL	NAN	SUS	PAT	POT	RAP	YOR
		R	8.13	10.65	14.39	12.05	9.23	NA	9.10	14.14	10.69	14.10	11.36	13.36	16.48	18.83	16.43	NA	NA	16.14	17.51	18.73	20.95	17.59
		F_{IS}	0.075	0.069	0.063	0.083	-0.033	0.018	0.071	0.045	0.028	0.041	0.137	0.036	0.001	0.044	0.038	0.119	0.104	0.048	0.059	0.070	0.091	-0.029
		H_E	0.605	0.723	0.797	0.804	0.772	0.718	0.753	0.858	0.835	0.859	0.815	0.873	0.904	0.922	0.862	0.943	0.909	0.861	0.911	0.898	0.919	0.884
		H_O	0.560	0.673	0.747	0.738	0.797	0.706	0.700	0.820	0.813	0.824	0.705	0.841	0.903	0.882	0.830	0.833	0.818	0.820	0.857	0.835	0.836	0.910
Af13	17	N	311	362	81	42	64	17	97	116	50	730	42	404	76	96	92	24	11	52	55	95	74	88
Range	152-186	N_a	7	10	9	7	7	6	8	9	8	11	6	11	13	12	11	10	6	8	9	12	10	12
		R	6.37	7.19	7.89	6.78	6.28	NA	7.71	7.92	7.54	7.61	5.74	7.51	10.40	9.62	9.78	NA	NA	7.74	8.58	10.14	8.59	9.10
		F_{IS}	0.218	0.116	0.096	0.203	0.128	0.141	0.151	0.017	0.381	0.157	-0.058	0.132	0.154	0.007	-0.035	0.015	-0.205	0.126	0.100	0.002	0.092	0.067
		H_E	0.694	0.793	0.791	0.775	0.770	0.818	0.837	0.842	0.772	0.809	0.721	0.772	0.854	0.849	0.840	0.888	0.762	0.836	0.868	0.865	0.863	0.853
		H_O	0.543	0.702	0.716	0.619	0.672	0.706	0.711	0.828	0.480	0.682	0.762	0.671	0.724	0.844	0.870	0.875	0.909	0.731	0.782	0.863	0.784	0.796
Af20	10	N	300	391	83	43	67	17	104	115	50	744	43	431	84	98	97	25	11	50	56	111	83	91
Range	159-177	N_a	3	3	4	3	4	3	3	3	3	3	2	4	3	4	3	3	3	6	4	4	4	5
		R	2.64	3.00	3.92	3.00	3.49	NA	2.99	2.64	2.89	2.62	2.00	3.33	3.00	3.56	2.99	NA	NA	4.64	3.56	3.28	3.39	3.93
		F_{IS}	-0.007	-0.034	0.171	0.271	0.058	-0.067	0.135	0.194	0.064	-0.089	-0.200	-0.100	-0.071	0.017	-0.114	0.034	0.167	-0.084	-0.151	-0.173	-0.026	0.082
		H_E	0.152	0.517	0.508	0.541	0.491	0.221	0.333	0.367	0.427	0.396	0.291	0.572	0.523	0.498	0.408	0.538	0.541	0.499	0.451	0.484	0.493	0.479
		H_O	0.153	0.535	0.422	0.395	0.463	0.235	0.289	0.296	0.400	0.432	0.349	0.629	0.560	0.490	0.454	0.520	0.455	0.540	0.518	0.568	0.506	0.440
Aps2A	18	N	315	386	83	42	66	17	93	112	40	718	41	411	85	99	96	23	12	50	59	116	83	92
Range	62-138	N_a	14	14	11	11	12	10	10	12	8	15	9	14	12	13	13	8	9	11	12	12	11	12
		R	8.20	8.39	8.34	10.69	10.78	NA	8.51	9.39	7.83	10.44	8.57	9.42	11.07	10.66	11.14	NA	NA	10.60	10.61	10.39	9.75	10.61
		F_{IS}	-0.033	0.012	-0.126	0.056	-0.027	0.041	0.013	-0.045	0.026	-0.004	0.263	0.109	0.275	0.142	0.155	0.065	0.369	0.313	0.165	0.273	0.129	0.202
		H_E	0.787	0.797	0.792	0.857	0.841	0.857	0.806	0.855	0.847	0.857	0.825	0.841	0.858	0.847	0.849	0.789	0.909	0.871	0.851	0.817	0.829	0.817
		H_O	0.813	0.788	0.892	0.810	0.864	0.824	0.796	0.893	0.825	0.861	0.610	0.749	0.624	0.727	0.719	0.739	0.583	0.600	0.712	0.595	0.723	0.652
AsaD042	22	N	283	342	83	40	66	17	89	117	50	732	45	400	84	100	97	24	12	46	55	104	69	93
Range	142-226	N_a	13	14	9	8	10	6	6	11	9	13	9	14	16	10	16	12	13	14	17	17	15	17
		R	8.84	9.87	7.96	7.62	7.94	NA	5.36	8.25	8.64	9.01	8.71	9.86	13.10	9.49	13.42	NA	NA	13.00	15.60	15.09	13.94	14.04
		F_{IS}	-0.046	-0.004	-0.009	-0.134	0.101	0.198	-0.085	-0.070	-0.088	-0.035	-0.030	-0.036	-0.079	0.054	-0.002	0.122	-0.069	0.105	0.024	0.002	0.029	0.015
		H_E	0.726	0.798	0.824	0.751	0.673	0.656	0.704	0.791	0.791	0.808	0.755	0.794	0.872	0.856	0.844	0.900	0.938	0.897	0.912	0.925	0.911	0.906
		H_O	0.760	0.801	0.831	0.850	0.606	0.529	0.764	0.846	0.860	0.836	0.778	0.823	0.941	0.810	0.845	0.792	1.000	0.804	0.891	0.923	0.884	0.893
AsaB020	15	N	293	363	85	43	66	17	91	113	50	732	45	392	78	95	95	22	12	53	58	110	75	95
Range	115-169	N_a	12	11	7	10	8	9	7	8	7	9	6	10	10	11	11	9	7	12	10	10	10	11
		R	7.48	8.12	6.77	9.60	7.80	NA	6.95	7.47	7.00	6.84	6.00	7.84	7.92	9.86	10.04	NA	NA	10.66	9.73	8.68	9.16	10.10
		F_{IS}	-0.015	-0.013	0.046	-0.095	0.129	0.142	0.092	-0.077	0.106	-0.033	-0.062	0.000	-0.013	0.052	-0.061	-0.024	0.116	0.014	0.066	-0.026	-0.025	-0.022
		H_E	0.774	0.835	0.826	0.829	0.816	0.888	0.750	0.838	0.827	0.804	0.775	0.791	0.785	0.855	0.854	0.844	0.750	0.822	0.830	0.806	0.820	0.814

Locus	A		STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD	DEL	NAN	SUS	PAT	POT	RAP	YOR
		H_O	0.785	0.846	0.788	0.907	0.712	0.765	0.681	0.903	0.740	0.831	0.822	0.791	0.795	0.811	0.905	0.864	0.667	0.811	0.776	0.827	0.840	0.832
AsaD429	15	N	295	363	84	42	64	17	94	113	40	710	45	389	80	97	92	24	12	49	56	113	72	93
Range	136-196	N_a	9	9	9	9	7	7	6	8	7	8	7	9	10	10	9	9	7	10	10	10	11	11
		R	7.66	8.83	8.04	8.74	6.86	NA	5.98	6.85	6.62	6.37	6.46	7.55	8.76	8.34	7.83	NA	NA	9.45	9.10	8.62	9.70	9.11
		F_{IS}	0.009	-0.015	-0.006	-0.001	-0.149	-0.201	-0.011	0.013	0.111	0.004	0.007	-0.019	-0.059	0.046	-0.006	-0.136	-0.005	0.095	0.217	0.156	0.142	0.039
		H_E	0.787	0.833	0.817	0.809	0.749	0.690	0.695	0.771	0.730	0.754	0.739	0.807	0.814	0.811	0.821	0.846	0.830	0.856	0.843	0.818	0.841	0.850
		H_O	0.780	0.846	0.821	0.810	0.859	0.824	0.702	0.761	0.650	0.751	0.733	0.823	0.863	0.773	0.826	0.958	0.833	0.776	0.661	0.690	0.722	0.817
AsD029	26	N	275	353	82	43	68	17	93	117	50	744	45	431	85	94	98	25	10	47	52	103	64	85
Range	148-264	N_a	13	12	8	8	9	7	8	10	8	11	8	15	12	11	13	11	7	11	13	15	11	13
		R	6.93	8.23	7.33	7.48	7.98	NA	6.52	8.12	7.74	7.59	7.33	8.30	9.48	9.29	10.88	NA	NA	10.24	11.67	11.42	9.69	11.01
		F_{IS}	0.013	-0.008	0.026	-0.034	-0.116	0.068	0.008	-0.024	-0.010	-0.038	-0.015	-0.033	0.017	-0.005	-0.021	-0.014	-0.191	0.094	0.017	0.001	0.018	-0.022
		H_E	0.789	0.702	0.676	0.607	0.659	0.756	0.759	0.776	0.733	0.744	0.745	0.703	0.813	0.784	0.800	0.829	0.763	0.751	0.861	0.826	0.843	0.818
		H_O	0.778	0.708	0.659	0.628	0.735	0.706	0.753	0.795	0.740	0.772	0.756	0.726	0.800	0.787	0.816	0.840	0.900	0.681	0.846	0.825	0.828	0.835
All loci		R	7.07	8.27	8.67	8.97	8.33	NA	7.37	8.14	7.64	8.51	7.10	8.63	9.99	10.47	10.93	NA	NA	10.11	10.36	11.02	10.65	11.03
		F_{IS}	0.031	0.018	0.024	0.014	0.036	0.024	0.038	0.000	0.032	0.011	0.011	0.019	0.010	0.033	-0.001	-0.007	0.057	0.078	0.040	0.041	0.050	0.008
		H_E	0.649	0.702	0.734	0.758	0.739	0.731	0.728	0.744	0.723	0.743	0.686	0.738	0.778	0.794	0.777	0.804	0.787	0.785	0.777	0.783	0.787	0.787
		H_O	0.629	0.689	0.717	0.747	0.713	0.714	0.700	0.744	0.700	0.735	0.679	0.724	0.771	0.768	0.778	0.810	0.744	0.724	0.746	0.752	0.748	0.783

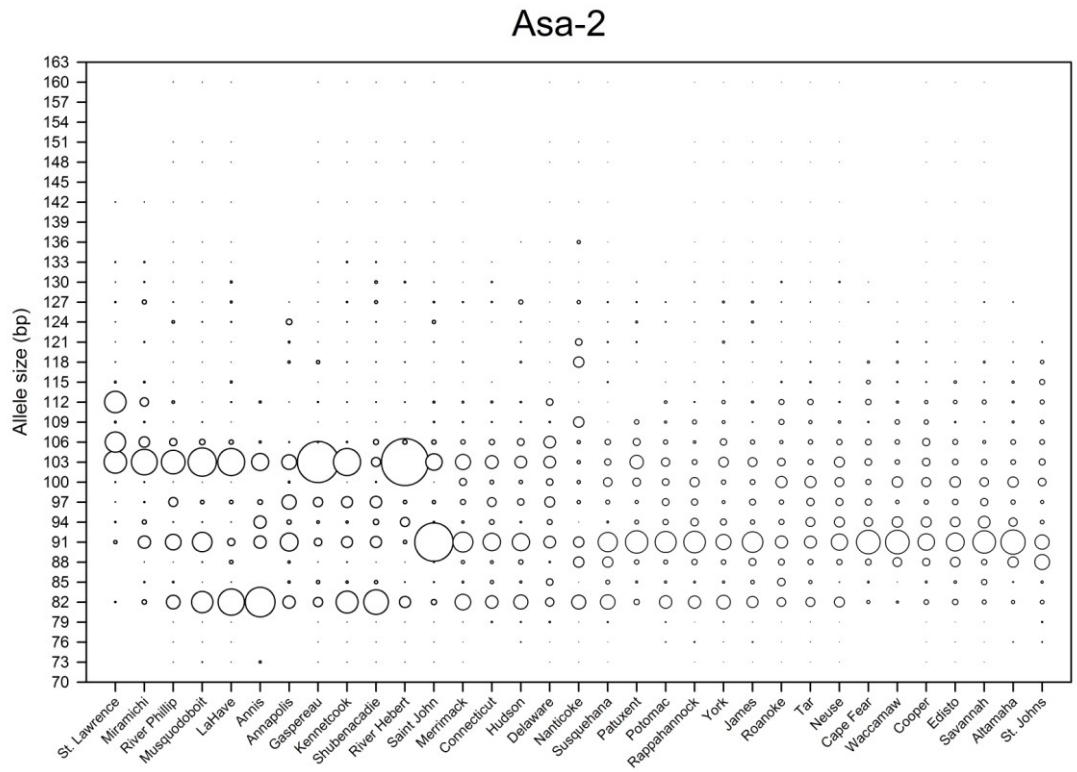
Appendix 5 *continued*

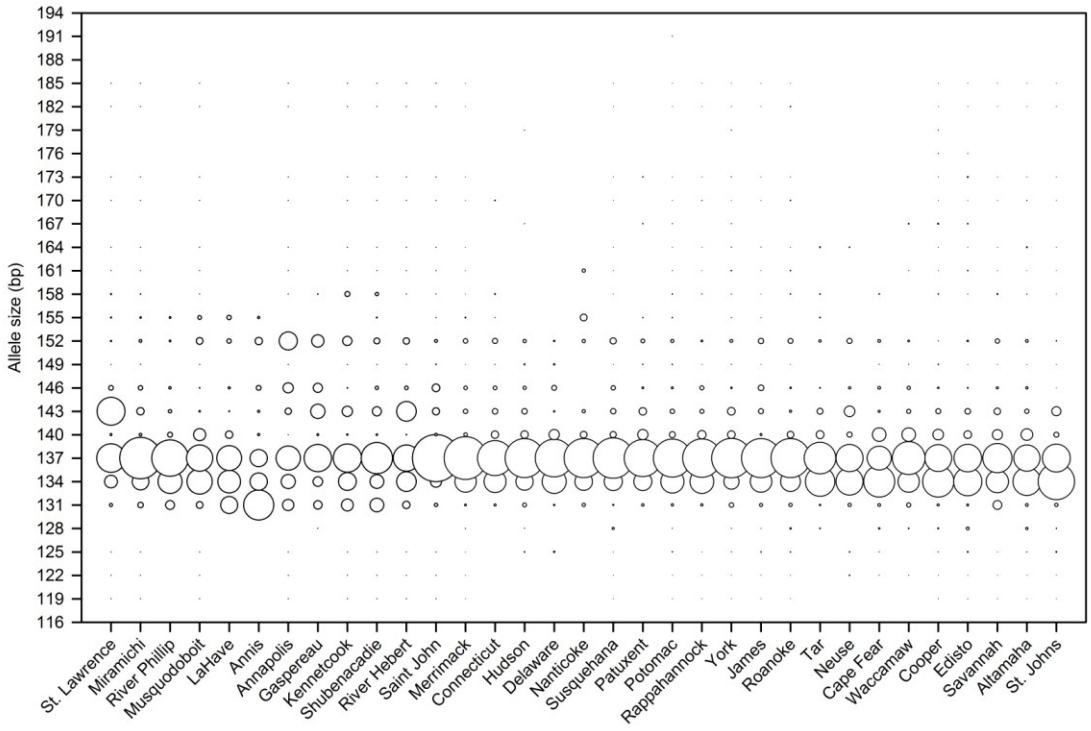
Locus		ROA	TAR	NEU	CF	WAC	COO	ED	SAV	ALT	STJ
Asa2	N	48	87	46	183	98	84	87	39	144	189
Range	N_a	16	19	15	16	16	16	15	14	17	18
	R	14.86	15.05	13.69	13.47	13.29	13.99	13.09	13.51	13.69	15.05
	F_{IS}	0.025	0.032	0.008	0.013	0.177	0.118	0.010	0.216	0.003	0.017
	H_E	0.919	0.914	0.898	0.875	0.867	0.903	0.894	0.880	0.864	0.904
	H_O	0.896	0.885	0.891	0.863	0.714	0.798	0.885	0.692	0.861	0.889
Asa4	N	49	97	47	173	100	90	94	38	161	182
Range	N_a	13	11	10	16	12	14	16	8	15	10
	R	11.12	8.86	9.07	9.78	9.04	9.90	11.42	7.85	9.96	6.84
	F_{IS}	0.046	-0.117	0.125	-0.061	0.013	-0.056	-0.066	-0.102	0.035	0.019
	H_E	0.727	0.730	0.777	0.758	0.760	0.747	0.769	0.789	0.773	0.695
	H_O	0.694	0.814	0.681	0.804	0.750	0.789	0.819	0.868	0.745	0.681
Asa8	N	49	82	47	169	95	85	66	33	155	186
Range	N_a	11	12	10	14	14	14	14	9	15	13
	R	10.11	9.52	9.27	9.64	11.18	11.20	10.94	9.00	11.52	8.82
	F_{IS}	0.071	0.162	0.205	0.130	0.135	0.087	0.065	-0.032	0.109	0.065
	H_E	0.791	0.814	0.801	0.802	0.827	0.811	0.793	0.764	0.818	0.782
	H_O	0.735	0.683	0.638	0.698	0.716	0.741	0.742	0.788	0.729	0.731
Asa16	N	49	95	47	180	98	90	88	37	156	171
Range	N_a	6	7	6	10	6	8	6	6	7	6
	R	5.89	6.55	5.91	6.33	5.50	6.37	5.67	5.89	5.33	4.12
	F_{IS}	-0.037	-0.085	-0.078	-0.082	0.017	0.002	-0.025	0.085	0.034	-0.108
	H_E	0.669	0.660	0.652	0.637	0.623	0.646	0.588	0.708	0.657	0.554
	H_O	0.694	0.716	0.702	0.689	0.612	0.644	0.602	0.649	0.635	0.614
Aa14	N	46	95	44	181	98	83	95	39	142	190
Range	N_a	17	25	15	26	25	25	28	19	31	28
	R	15.22	17.36	13.04	15.78	16.66	17.83	19.54	17.52	17.39	17.44
	F_{IS}	-0.062	0.017	-0.042	-0.018	-0.081	-0.011	0.062	0.031	0.037	0.003
	H_E	0.779	0.696	0.655	0.684	0.727	0.715	0.774	0.767	0.694	0.766
	H_O	0.826	0.684	0.682	0.696	0.786	0.723	0.726	0.744	0.669	0.763
Af6	N	48	94	44	181	95	68	82	34	106	169
Range	N_a	17	28	21	22	26	26	25	21	28	22

Locus		ROA	TAR	NEU	CF	WAC	COO	ED	SAV	ALT	STJ
	R	16.07	20.91	19.90	16.84	20.18	21.16	20.63	20.91	21.72	16.20
	F_{IS}	0.115	0.139	0.079	0.029	0.049	0.075	0.023	0.060	0.032	0.067
	H_E	0.917	0.938	0.937	0.893	0.941	0.937	0.936	0.938	0.945	0.913
	H_O	0.813	0.809	0.864	0.867	0.895	0.868	0.915	0.882	0.915	0.852
Af13	N	48	95	45	173	99	86	80	37	136	184
Range	N_a	11	11	9	11	13	11	15	11	14	14
	R	9.94	9.76	8.72	9.76	10.32	9.36	11.80	10.57	10.99	10.45
	F_{IS}	-0.064	0.046	-0.027	0.017	0.111	0.121	0.017	0.047	0.093	0.030
	H_E	0.843	0.860	0.866	0.864	0.840	0.846	0.864	0.850	0.859	0.852
	H_O	0.896	0.821	0.889	0.850	0.748	0.744	0.850	0.811	0.779	0.826
Af20	N	49	81	47	185	83	75	92	39	160	189
Range	N_a	4	4	5	8	4	5	5	3	5	4
	R	3.89	3.67	4.59	4.46	3.74	4.24	4.18	3.00	4.26	3.65
	F_{IS}	0.015	-0.163	0.122	-0.131	-0.054	0.035	-0.024	-0.134	-0.061	0.017
	H_E	0.435	0.425	0.412	0.473	0.457	0.401	0.499	0.408	0.471	0.506
	H_O	0.429	0.494	0.362	0.535	0.482	0.387	0.511	0.462	0.500	0.497
Aps2A	N	48	94	47	174	98	82	86	38	158	188
Range	N_a	11	12	11	14	11	9	12	10	12	11
	R	10.65	10.37	10.44	10.84	9.90	8.36	10.37	9.85	10.24	9.14
	F_{IS}	0.312	0.197	0.211	0.088	0.247	0.152	0.153	0.132	0.122	0.170
	H_E	0.845	0.821	0.834	0.787	0.839	0.819	0.850	0.817	0.843	0.807
	H_O	0.583	0.660	0.660	0.718	0.633	0.695	0.721	0.711	0.741	0.670
AsaD042	N	49	90	47	167	98	87	78	38	131	188
Range	N_a	16	16	17	18	17	16	16	16	20	18
	R	14.80	13.67	15.44	14.12	13.99	13.42	13.91	15.57	14.95	14.32
	F_{IS}	-0.043	-0.011	0.007	0.050	-0.046	0.009	0.026	0.054	0.015	0.068
	H_E	0.920	0.902	0.921	0.907	0.897	0.893	0.909	0.918	0.922	0.901
	H_O	0.959	0.911	0.915	0.862	0.939	0.885	0.885	0.868	0.908	0.840
AsaB020	N	49	96	47	178	99	92	93	39	159	190
Range	N_a	10	10	11	11	11	9	12	8	12	10
	R	9.21	9.22	9.64	8.66	8.69	8.56	8.82	7.82	7.96	6.96
	F_{IS}	-0.082	-0.012	0.044	-0.040	0.120	0.105	-0.003	0.066	-0.029	-0.063
	H_E	0.812	0.772	0.801	0.800	0.757	0.777	0.740	0.741	0.776	0.777

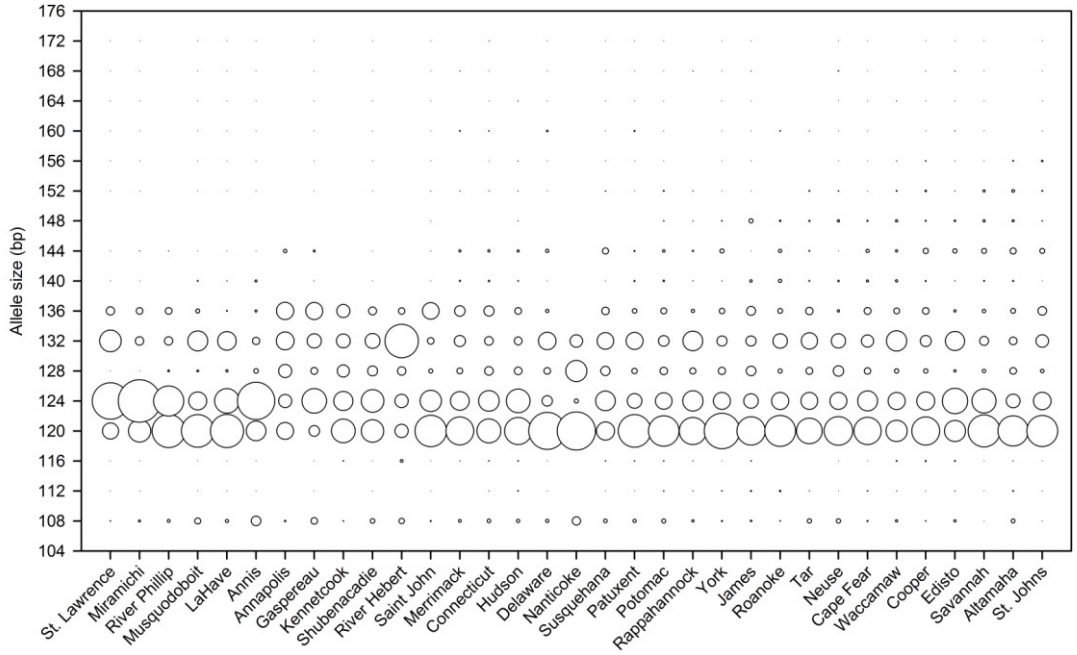
Locus		ROA	TAR	NEU	CF	WAC	COO	ED	SAV	ALT	STJ
	H_O	0.878	0.781	0.766	0.832	0.667	0.696	0.742	0.692	0.799	0.826
AsaD429	N	45	90	45	184	96	83	83	37	158	182
Range	N_a	11	11	9	12	12	10	11	9	11	10
	R	10.43	8.54	8.73	9.20	9.61	9.11	9.43	8.99	9.89	8.75
	F_{IS}	0.256	0.038	0.239	0.038	0.125	0.030	-0.027	-0.028	0.094	-0.028
	H_E	0.863	0.831	0.845	0.847	0.809	0.832	0.810	0.842	0.824	0.668
	H_O	0.644	0.800	0.644	0.815	0.708	0.807	0.831	0.865	0.747	0.687
AsD029	N	47	91	44	160	94	80	76	38	129	188
Range	N_a	13	13	13	16	14	14	14	11	17	13
	R	11.70	11.07	12.12	11.39	11.87	11.48	12.23	10.85	12.89	10.44
	F_{IS}	0.005	0.028	0.020	-0.012	0.050	-0.013	-0.001	-0.094	0.053	0.014
	H_E	0.856	0.803	0.881	0.852	0.873	0.864	0.854	0.867	0.875	0.858
	H_O	0.851	0.780	0.864	0.863	0.830	0.875	0.855	0.947	0.830	0.846
All loci	R	11.07	11.12	10.81	10.79	11.08	11.15	11.69	10.87	11.60	10.17
	F_{IS}	0.047	0.032	0.071	0.009	0.073	0.053	0.019	0.030	0.045	0.026
	H_E	0.798	0.782	0.791	0.783	0.786	0.784	0.791	0.791	0.794	0.768
	H_O	0.761	0.757	0.735	0.776	0.729	0.742	0.776	0.768	0.758	0.748

Appendix 6 Allele frequency distributions of 13 microsatellite loci for 33 shad populations examined in this study.

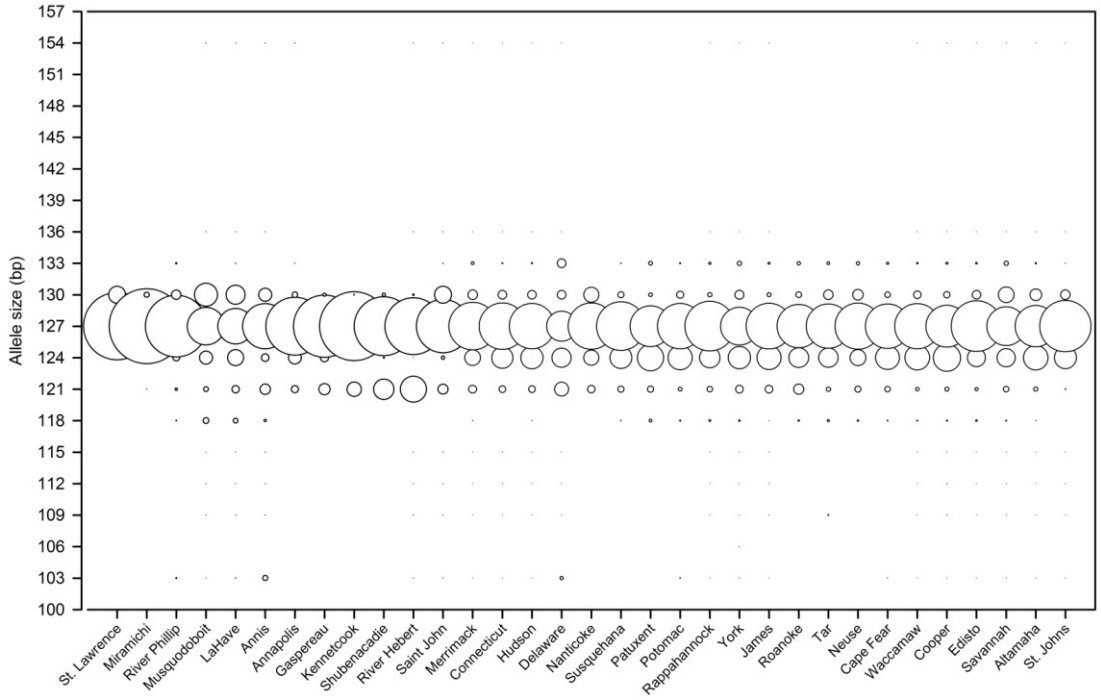

Asa-2
Allele size (bp)
163
160
157
154
151
148
145
142
139
136
133
130
127
124
121
118
115
112
109
106
103
100
97
94
91
88
85
82
79
76
73
70
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns


Asa-4
Allele size (bp)
194
191
188
185
182
179
176
173
170
167
164
161
158
155
152
149
146
143
140
137
134
131
128
125
122
119
116
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

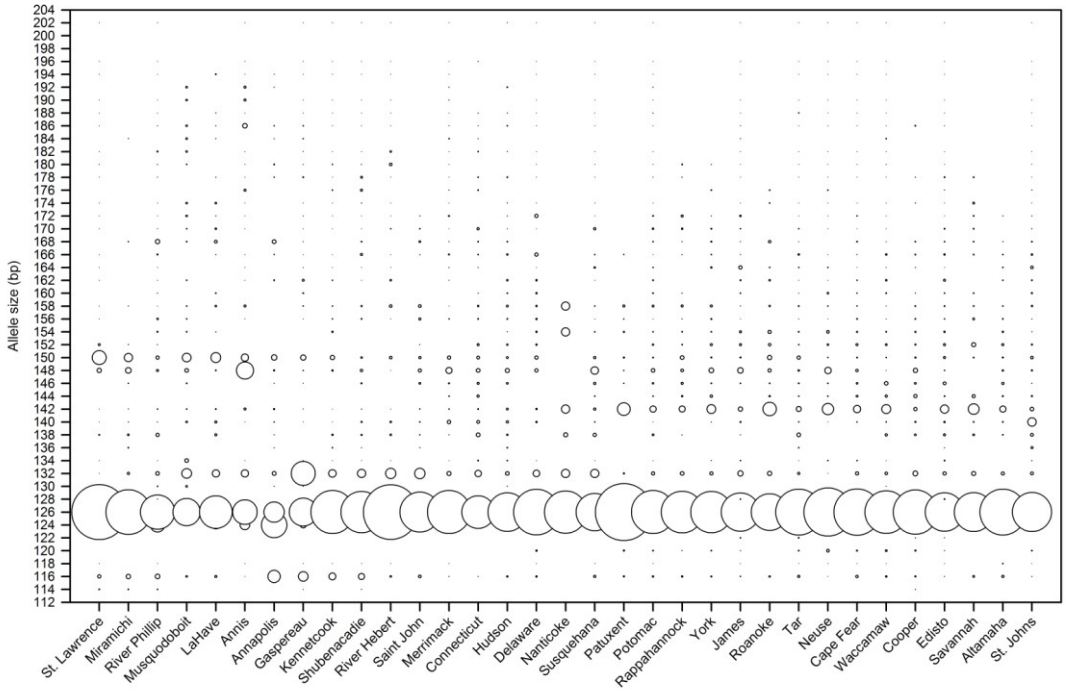
Asa-8


Allele size (bp)
176
172
168
164
160
156
152
148
144
140
136
132
128
124
120
116
112
108
104
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

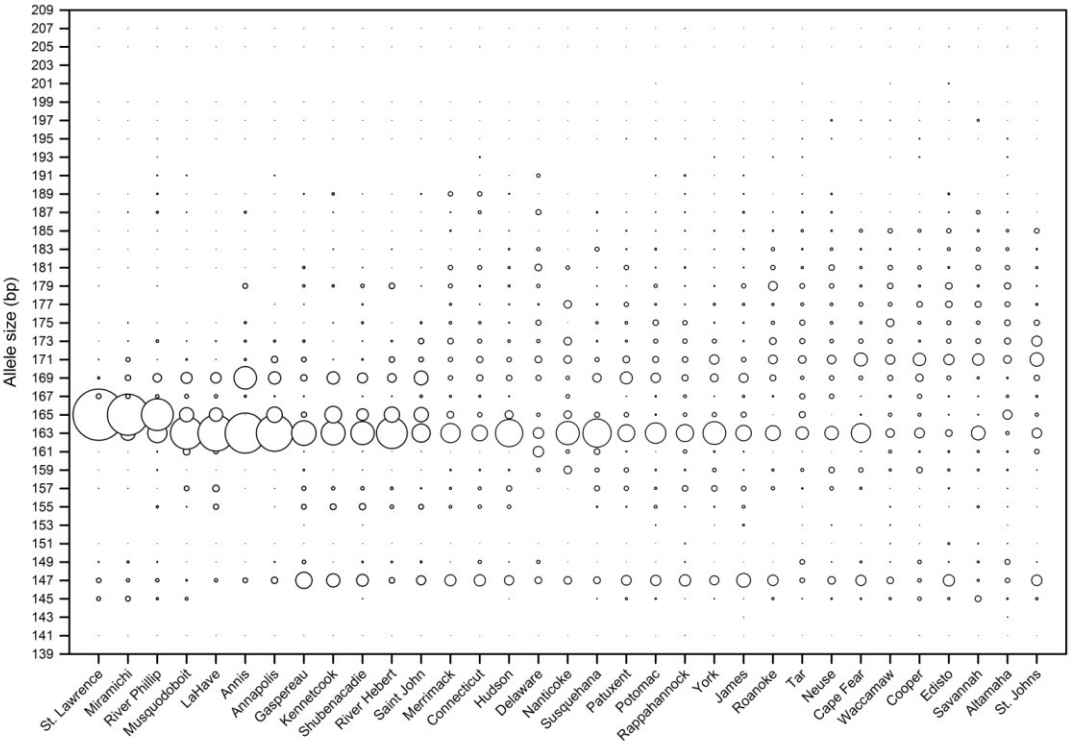
Asa-16

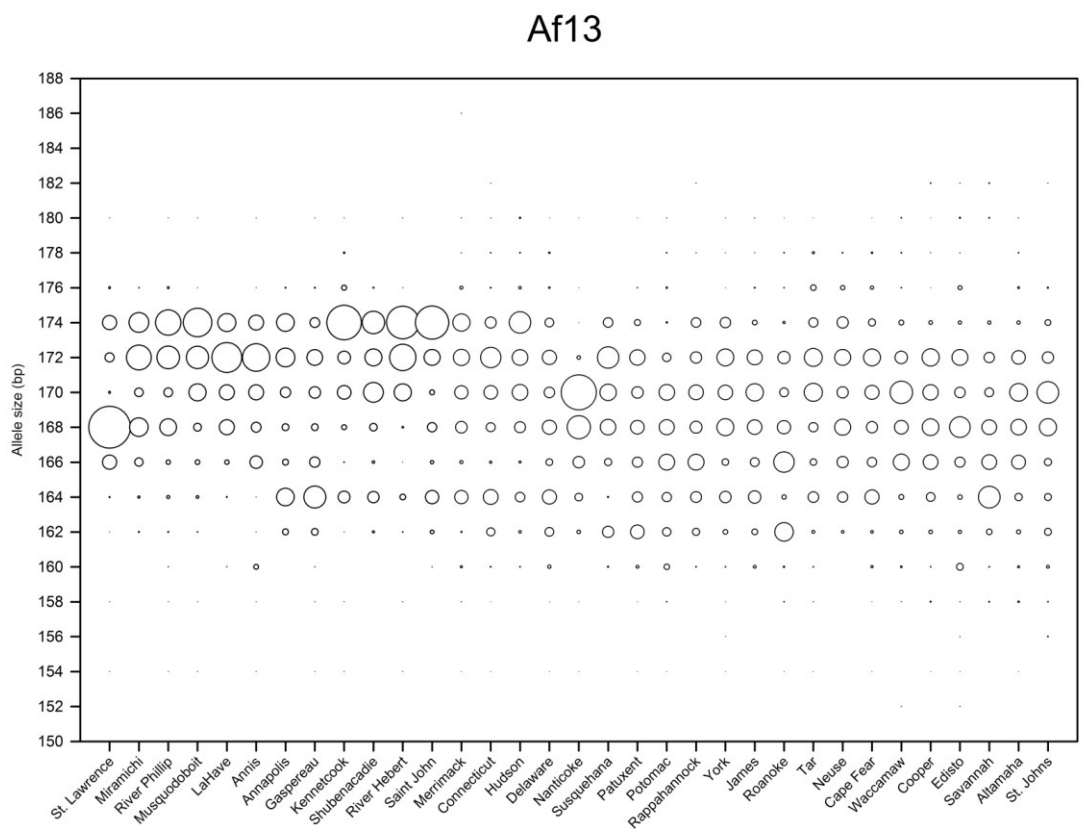

Allele size (bp)
157
154
151
148
145
142
139
136
133
130
127
124
121
118
115
112
109
106
103
100
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

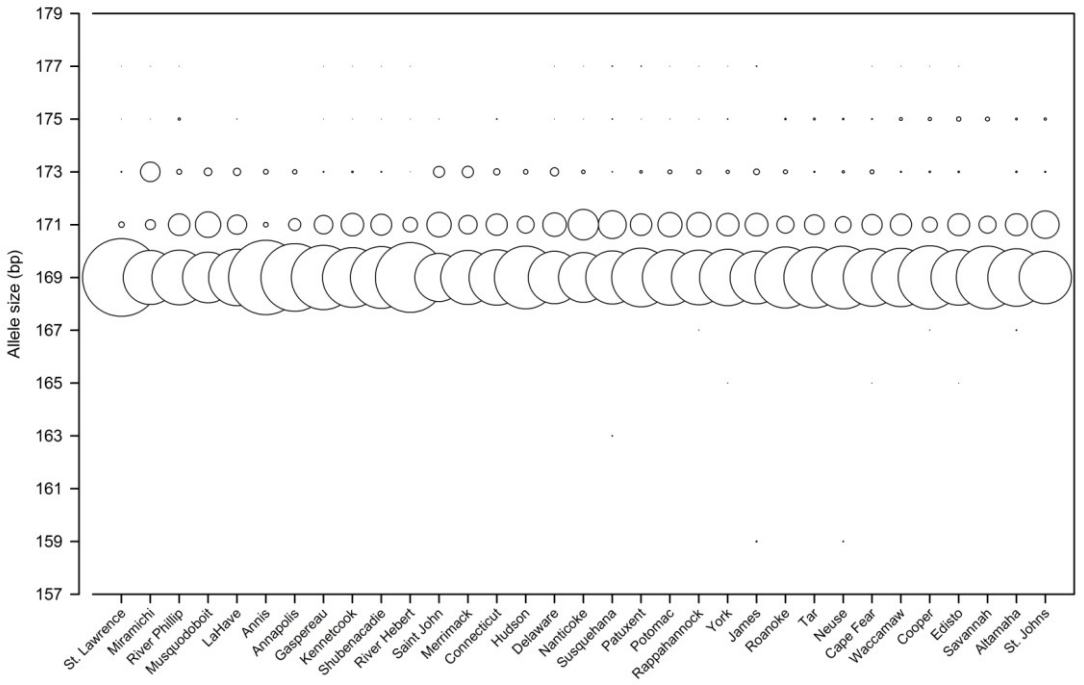
Aa-14


Allele size (bp)
204
202
200
198
196
194
192
190
188
186
184
182
180
178
176
174
172
170
168
166
164
162
160
158
156
154
152
150
148
146
144
142
140
138
136
134
132
130
128
126
124
122
120
118
116
114
112
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

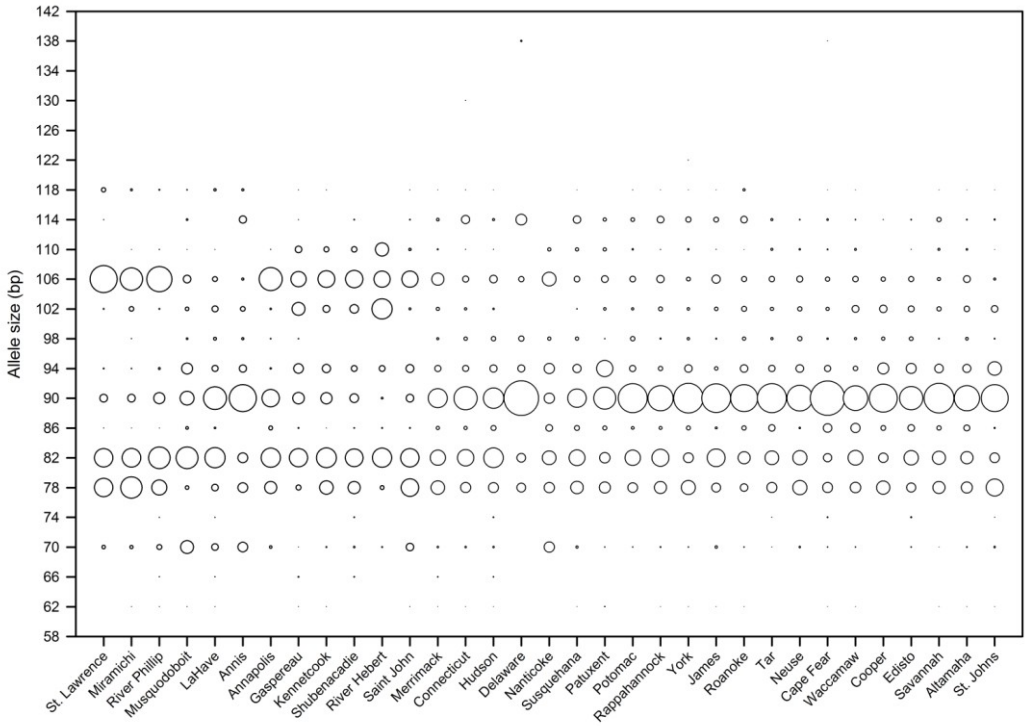
Af6


Allele size (bp)
209
207
205
203
201
199
197
195
193
191
189
187
185
183
181
179
177
175
173
171
169
167
165
163
161
159
157
155
153
151
149
147
145
143
141
139
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

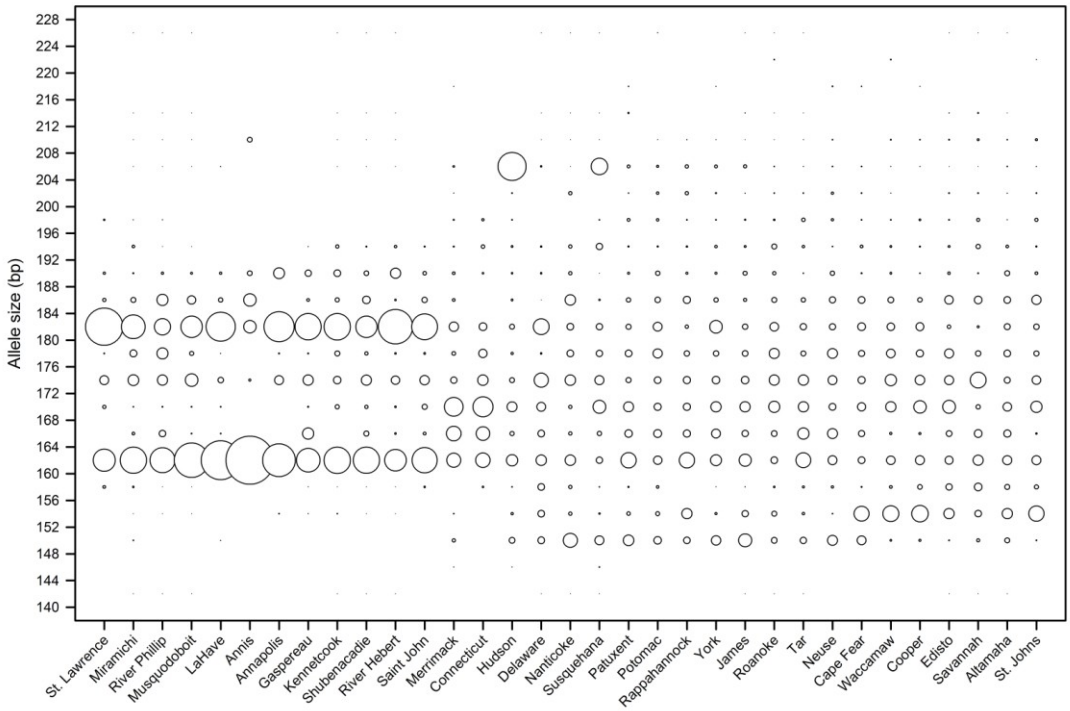

Af13
Allele size (bp)
188
186
184
182
180
178
176
174
172
170
168
166
164
162
160
158
156
154
152
150
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

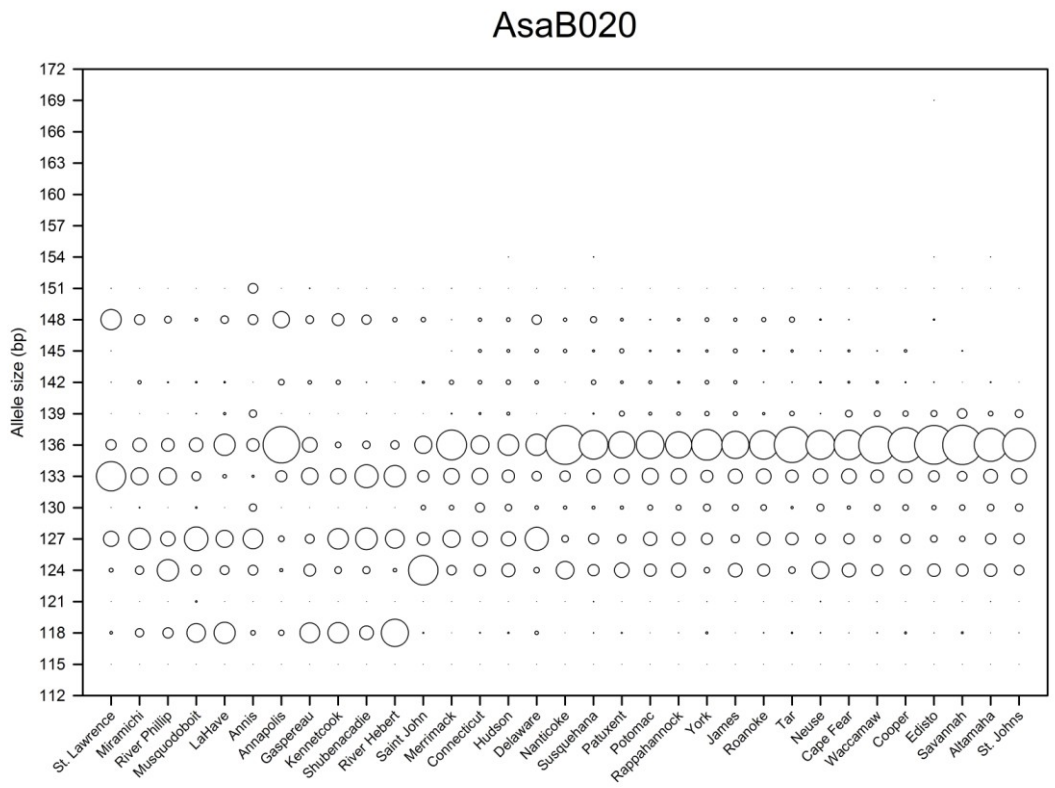

Af20
Allele size (bp)
179
177
175
173
171
169
167
165
163
161
159
157
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

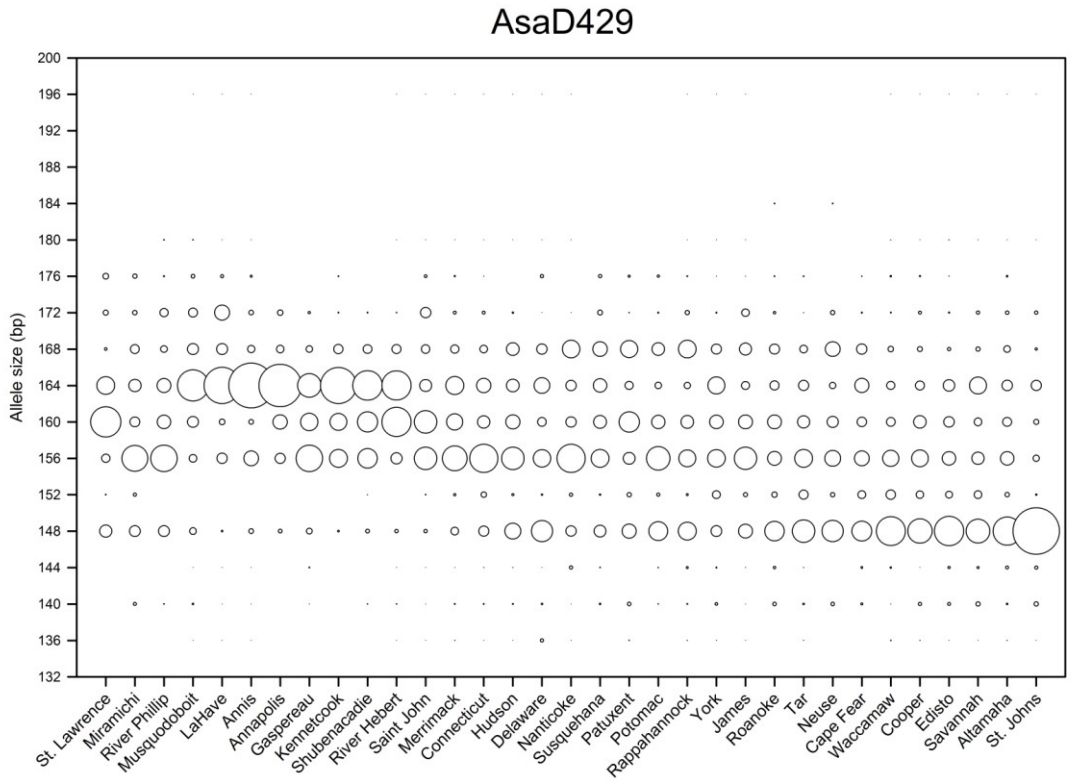
Aps2A

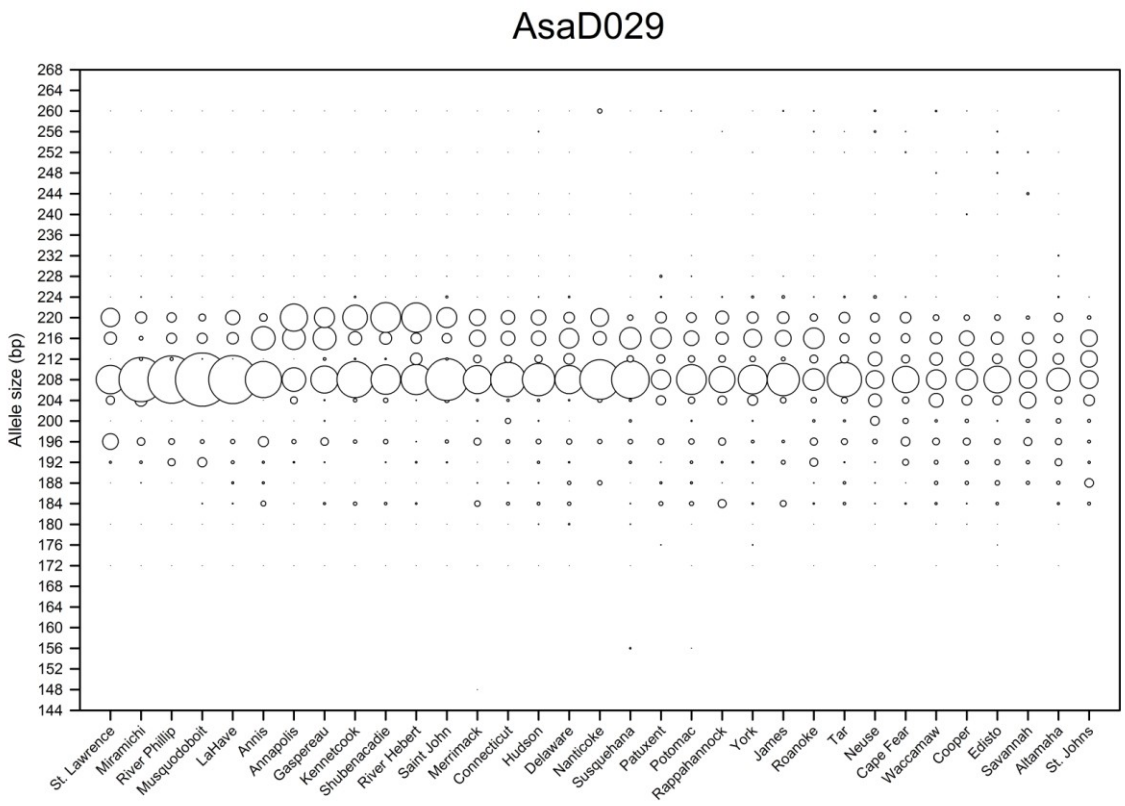


AsaD042



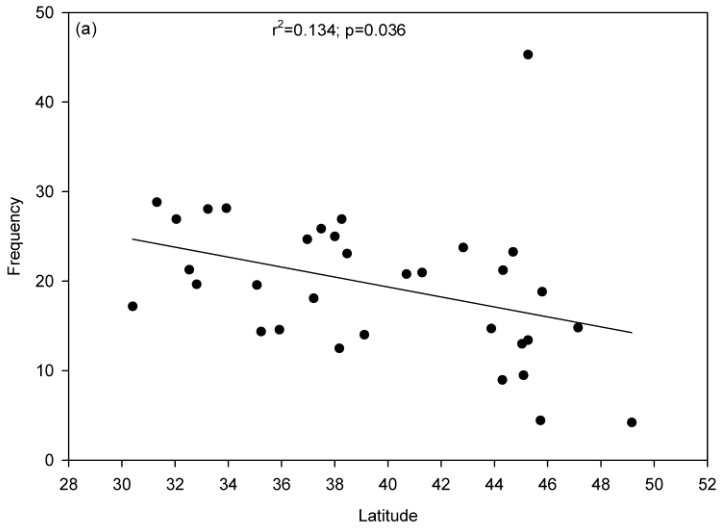

AsaB020
Allele size (bp)
172
169
166
163
160
157
154
151
148
145
142
139
136
133
130
127
124
121
118
115
112
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns


AsaD429
Allele size (bp)
200
196
192
188
184
180
176
172
168
164
160
156
152
148
144
140
136
132
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

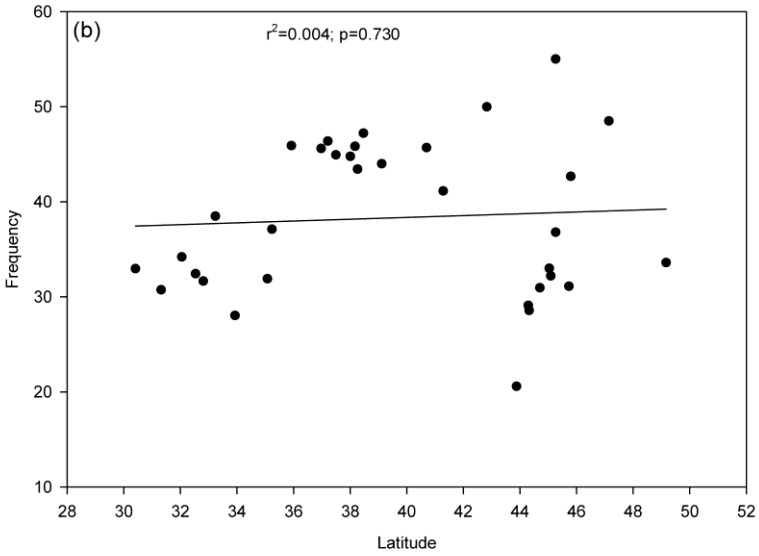

AsaD029
Allele size (bp)
268
264
260
256
252
248
244
240
236
232
228
224
220
216
212
208
204
200
196
192
188
184
180
176
172
168
164
160
156
152
148
144
St. Lawrence
Miramichi
River Phillip
Musquodoboit
LaHave
Annis
Annapolis
Gaspereau
Kennetcook
Shubenacadie
River Hebert
Saint John
Merrimack
Connecticut
Hudson
Delaware
Nanticoke
Susquehana
Patuxent
Potomac
Rappahannock
York
James
Roanoke
Tar
Neuse
Cape Fear
Waccamaw
Cooper
Edisto
Savannah
Altamaha
St. Johns

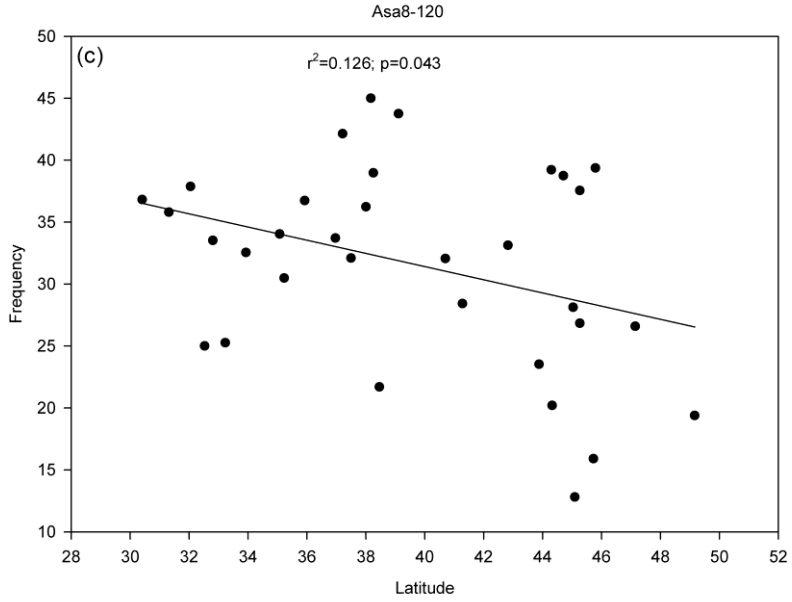
Appendix 7 Clinal variation of the most common allele for each locus examined in this study

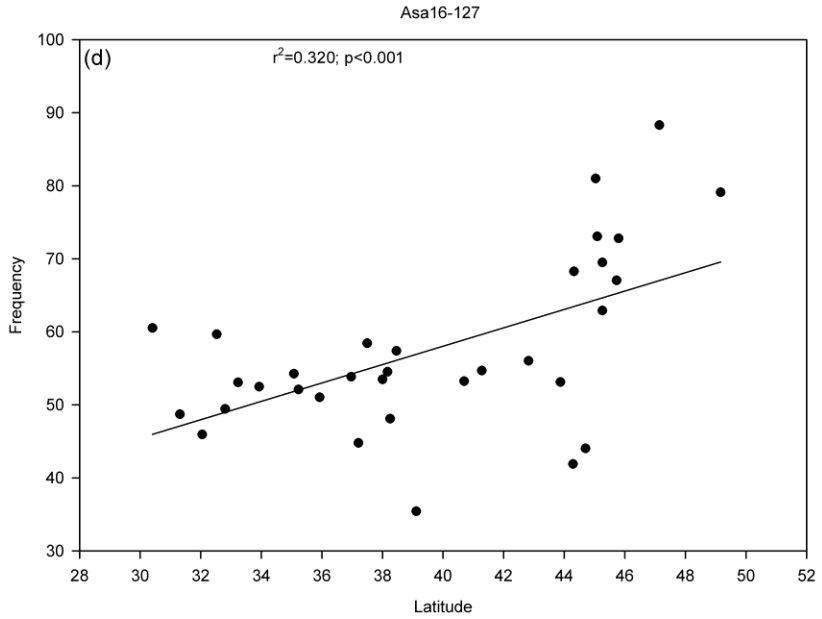
Asa2-91

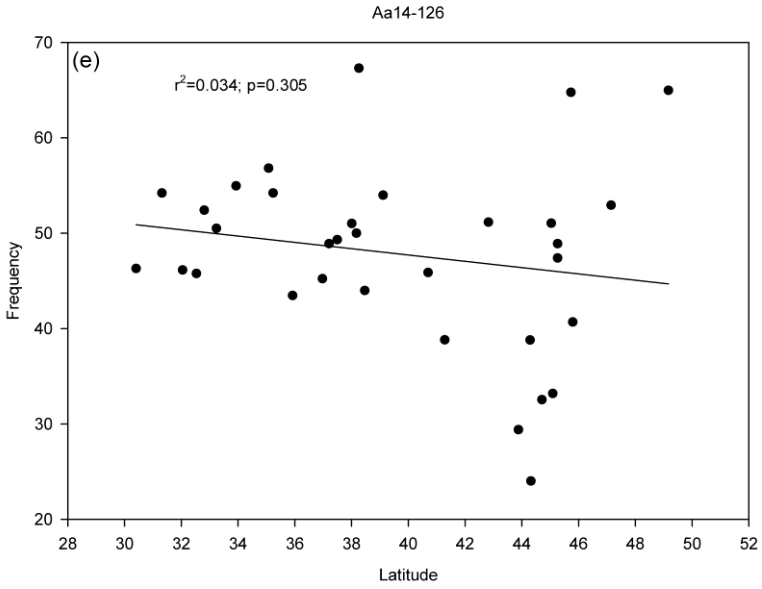


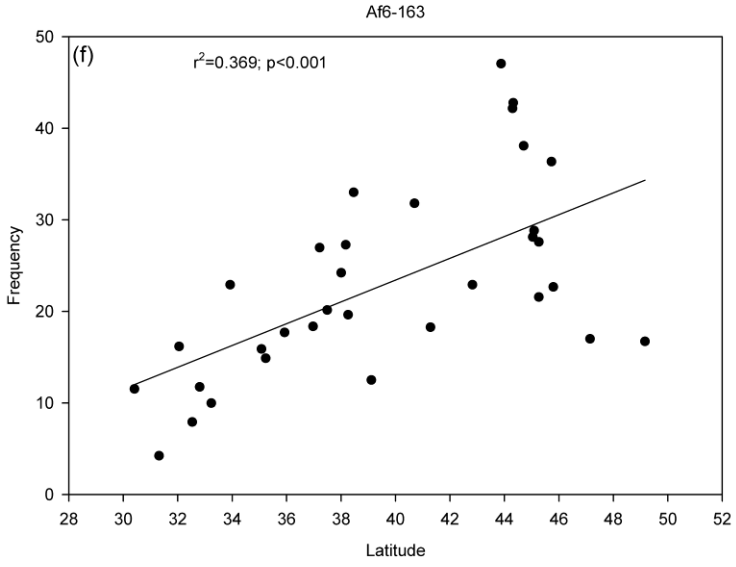
Asa4-137

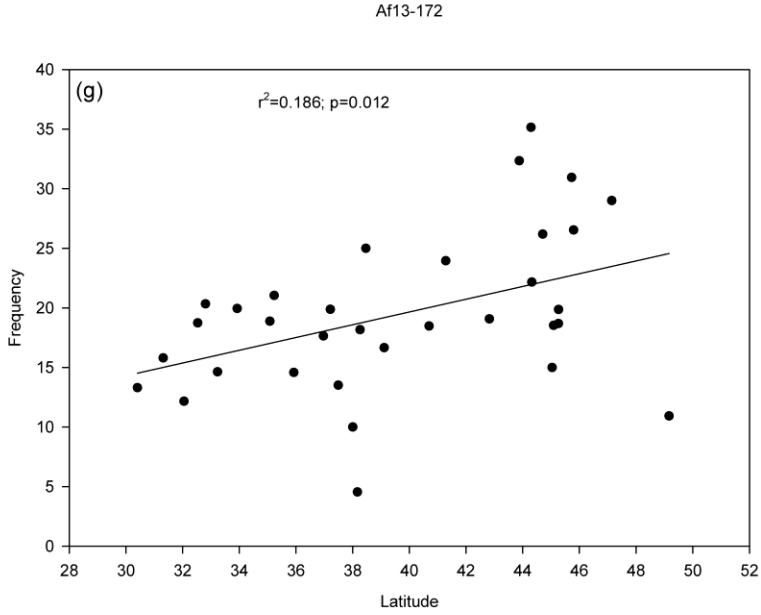


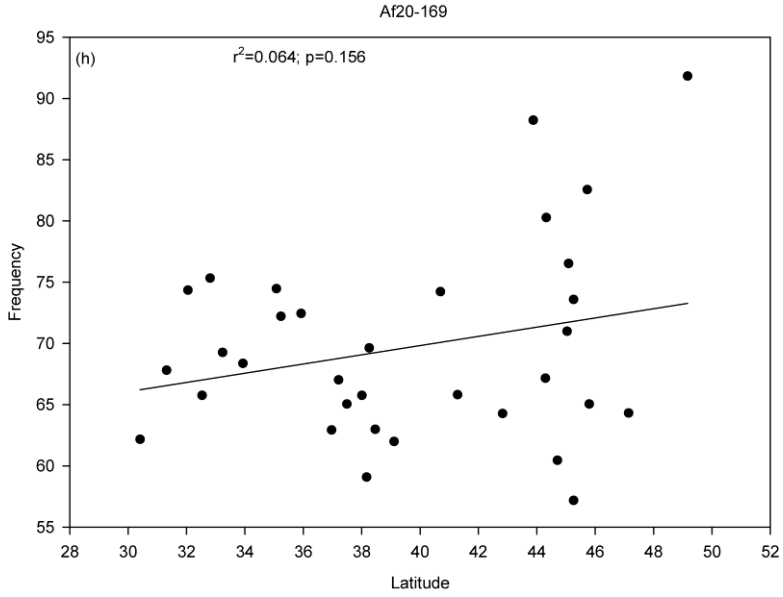

Asa8-120
(c)
r^2=0.126; p=0.043
Frequency
Latitude
50
45
40
35
30
25
20
15
10
28
30
32
34
36
38
40
42
44
46
48
50
52

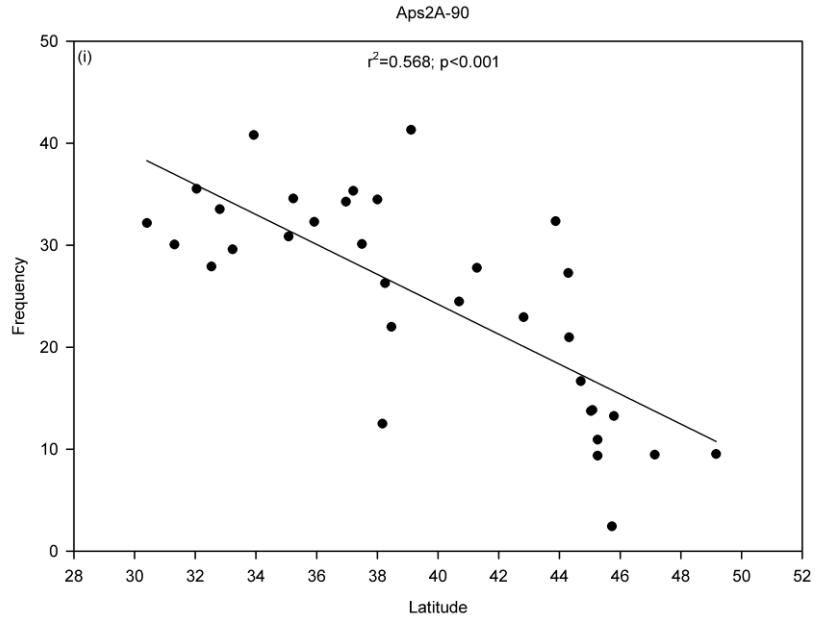

Asa16-127
(d)
r^2=0.320; p<0.001
Frequency
Latitude
100
90
80
70
60
50
40
30
28
30
32
34
36
38
40
42
44
46
48
50
52

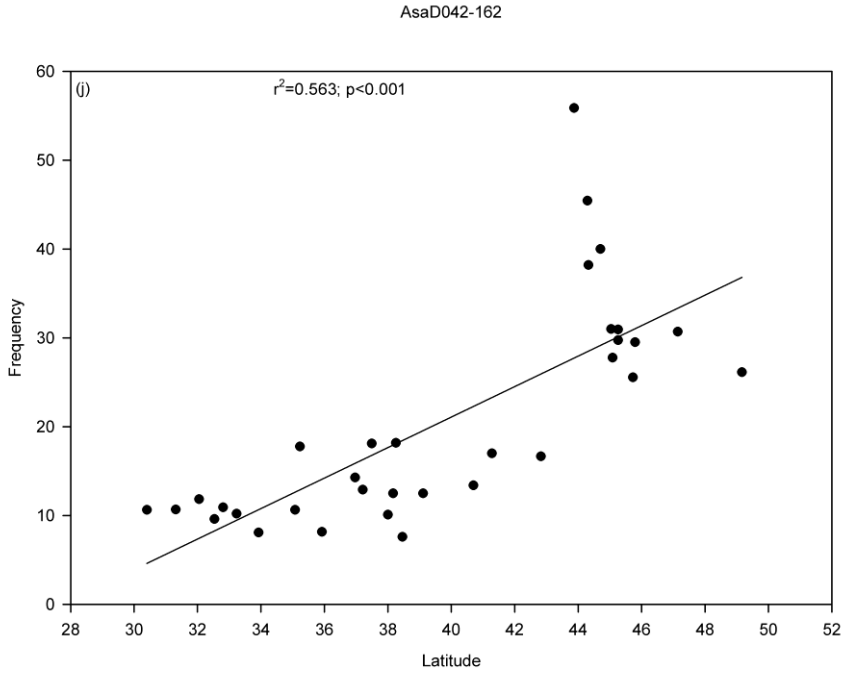

Aa14-126
(e)
r^2=0.034; p=0.305
Frequency
Latitude
70
60
50
40
30
20
28
30
32
34
36
38
40
42
44
46
48
50
52

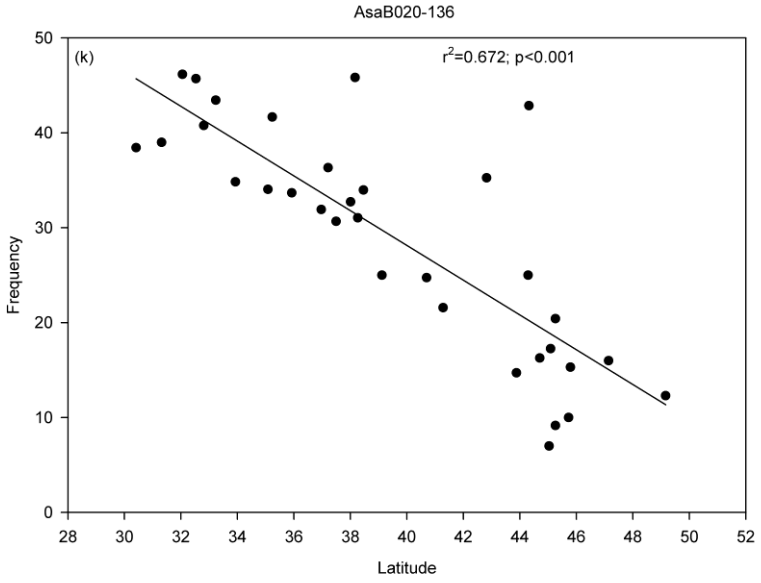

Af6-163
(f)
r^2=0.369; p<0.001
Frequency
Latitude
50
40
30
20
10
0
28
30
32
34
36
38
40
42
44
46
48
50
52

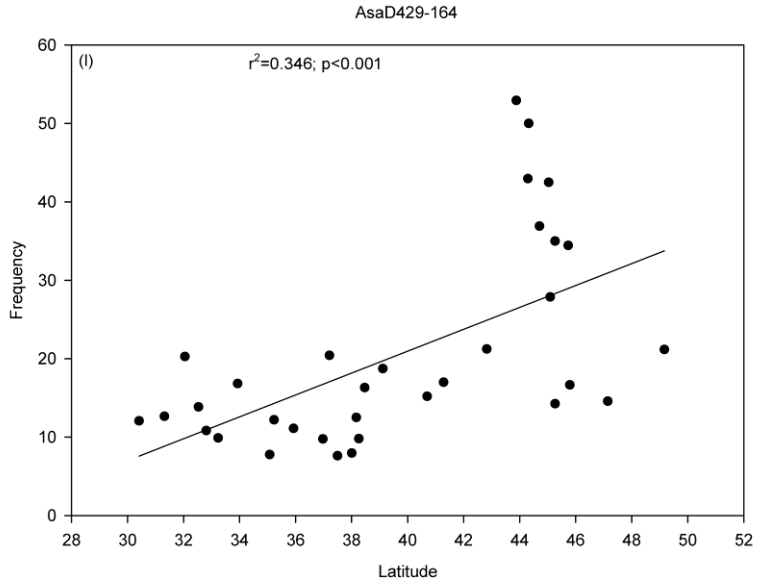

Af13-172
(g)
r^2=0.186; p=0.012
Frequency
Latitude
0
5
10
15
20
25
30
35
40
28
30
32
34
36
38
40
42
44
46
48
50
52

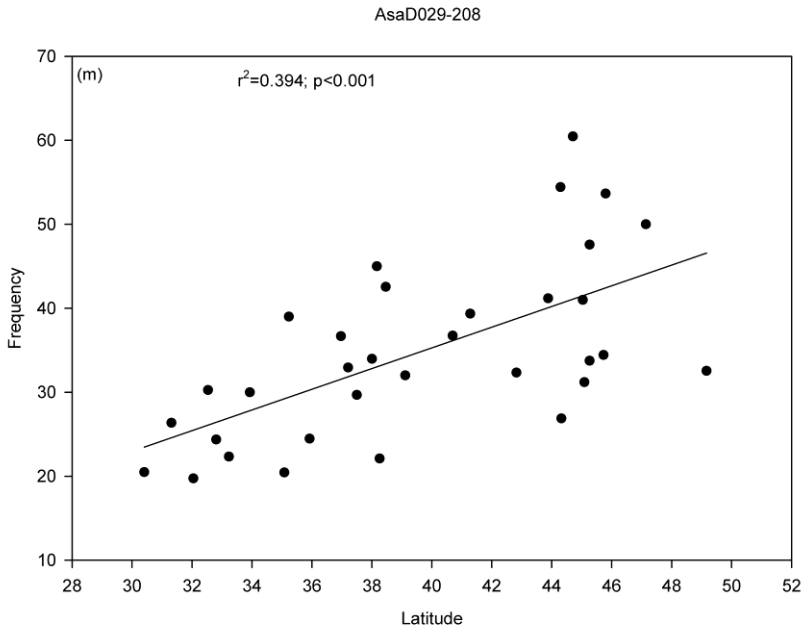

Af20-169
(h)
r^2=0.064; p=0.156
Frequency
Latitude
55
60
65
70
75
80
85
90
95
28
30
32
34
36
38
40
42
44
46
48
50
52


Aps2A-90
(i)
r²=0.568; p<0.001
Frequency
Latitude
0
10
20
30
40
50
28
30
32
34
36
38
40
42
44
46
48
50
52


AsaD042-162
(j)
r²=0.563; p<0.001
Frequency
Latitude
0
10
20
30
40
50
60
28
30
32
34
36
38
40
42
44
46
48
50
52


AsaB020-136
(k)
r^2=0.672; p<0.001
Frequency
Latitude
0
10
20
30
40
50
28
30
32
34
36
38
40
42
44
46
48
50
52


AsaD429-164
(l)
r^2=0.346; p<0.001
Frequency
Latitude
0
10
20
30
40
50
60
28
30
32
34
36
38
40
42
44
46
48
50
52


AsaD029-208
(m)
r²=0.394; p<0.001
Frequency
Latitude
70
60
50
40
30
20
10
28
30
32
34
36
38
40
42
44
46
48
50
52

Appendix 8 (back pocket) Genetic differentiation among 33 spawning runs of American shad from across the species' range. Matrix of pairwise F_{ST} estimates (θ; Weir and Cockerham 1984) below diagonal, and standardized F'_{ST} estimates (Hedrick 2005) above diagonal (bold values indicates non-significant differentiation; $p>0.05$).

Appendix 8

	STL	MR	RP	MUS	LH	AR	ANN	GR	KR	SHU	RH	SJR	MER	CON	HUD	DEL	NAN	SUS	PAT	POT	RAP	YOR
STL		0.1224	0.1818	0.3068	0.2955	0.3322	0.2720	0.2500	0.2122	0.2173	0.2268	0.2581	0.2725	0.3092	0.2853	0.3222	0.3275	0.3125	0.3002	0.3161	0.3011	0.3031
MR	0.0396		0.0470	0.1868	0.1970	0.2401	0.2257	0.1849	0.1276	0.1509	0.2081	0.1347	0.1567	0.1811	0.1801	0.2391	0.2223	0.2021	0.2351	0.2184	0.1959	0.2155
RP	0.0574	0.0134		0.0902	0.1186	0.1768	0.1722	0.1456	0.0798	0.1026	0.1899	0.0759	0.0980	0.1098	0.1128	0.1720	0.1283	0.1320	0.1784	0.1486	0.1338	0.1519
MUS	0.0949	0.0517	0.0230		0.0079	0.0973	0.1788	0.1738	0.1040	0.1318	0.1697	0.1361	0.1392	0.1427	0.1412	0.1510	0.1473	0.1456	0.1997	0.1855	0.1699	0.1512
LH	0.0930	0.0558	0.0312	**0.0020**		0.0554	0.1492	0.1718	0.1133	0.1348	0.1759	0.1848	0.1545	0.1649	0.1575	0.1670	0.1710	0.1657	0.2156	0.2023	0.1924	0.1497
AR	0.1070	0.0691	0.0472	0.0247	0.0146		0.1403	0.1763	0.1416	0.1355	0.2406	0.2401	0.1823	0.1782	0.1747	0.2204	0.2425	0.1897	0.2439	0.2186	0.2069	0.1805
ANN	0.0864	0.0649	0.0463	0.0463	0.0398	0.0380		0.1180	0.1003	0.1232	0.1931	0.1896	0.1567	0.1838	0.1846	0.2322	0.2054	0.1827	0.2335	0.2230	0.2050	0.1863
GR	0.0774	0.0518	0.0379	0.0434	0.0443	0.0460	0.0311		0.0698	0.0925	0.1047	0.1670	0.1378	0.1341	0.1640	0.2047	0.1987	0.1741	0.2008	0.1922	0.1844	0.1849
KR	0.0684	0.0370	0.0216	0.0270	0.0304	0.0387	0.0275	0.0185		0.0147	0.0586	0.1158	0.1159	0.1394	0.1397	0.1922	0.1835	0.1699	0.2039	0.1998	0.1770	0.1730
SHU	0.0647	0.0416	0.0267	0.0331	0.0348	0.0354	0.0324	0.0237	0.0039		0.0913	0.1289	0.1021	0.1212	0.1247	0.1672	0.1430	0.1518	0.1771	0.1702	0.1448	0.1475
RH	0.0764	0.0633	0.0546	0.0472	0.0502	0.0709	0.0561	0.0294	0.0173	0.0254		0.2150	0.2060	0.2383	0.2263	0.2640	0.2852	0.2564	0.2523	0.2895	0.2657	0.2550
SJR	0.0785	0.0377	0.0200	0.0346	0.0483	0.0635	0.0505	0.0433	0.0310	0.0334	0.0606		0.0921	0.1226	0.1275	0.1794	0.1386	0.1481	0.1634	0.1615	0.1437	0.1519
MER	0.0810	0.0418	0.0239	0.0322	0.0372	0.0440	0.0388	0.0330	0.0287	0.0248	0.0544	0.0226		0.0075	0.0306	0.0390	0.0371	0.0366	0.0518	0.0356	0.0317	0.0250
CON	0.0893	0.0469	0.0258	0.0317	0.0382	0.0412	0.0439	0.0310	0.0332	0.0286	0.0606	0.0292	**0.0016**		0.0388	0.0328	0.0548	0.0338	0.0675	0.0336	0.0368	0.0332
HUD	0.0846	0.0480	0.0275	0.0327	0.0379	0.0422	0.0457	0.0393	0.0346	0.0303	0.0597	0.0313	0.0068	0.0083		0.0644	0.0548	0.0192	0.0638	0.0474	0.0417	0.0397
DEL	0.0945	0.0620	0.0405	0.0333	0.0387	0.0508	0.0557	0.0472	0.0461	0.0390	0.0686	0.0424	0.0082	**0.0066**	**0.0136**		0.0404	0.0511	0.0272	0.0174	0.0243	0.0059
NAN	0.0992	0.0595	0.0314	0.0338	0.0412	0.0586	0.0511	0.0475	0.0459	0.0344	0.0778	0.0338	**0.0081**	0.0114	**0.0120**	**0.0082**		0.0070	0.0285	-0.0047	0.0089	0.0076
SUS	0.0929	0.0536	0.0319	0.0332	0.0395	0.0451	0.0449	0.0413	0.0417	0.0365	0.0675	0.0360	0.0080	0.0071	**0.0042**	**0.0105**	**0.0015**		0.0458	0.0269	0.0229	0.0268
PAT	0.0900	0.0630	0.0437	0.0463	0.0521	0.0590	0.0581	0.0483	0.0508	0.0431	0.0673	0.0402	0.0115	0.0144	0.0142	**0.0057**	**0.0062**	**0.0100**		0.0169	0.0133	0.0176
POT	0.0926	0.0575	0.0357	0.0423	0.0480	0.0519	0.0544	0.0454	0.0487	0.0409	0.0752	0.0392	0.0078	0.0071	0.0104	**0.0036**	**-0.0010**	0.0058	**0.0037**		-0.0033	0.0141
RAP	0.0886	0.0516	0.0321	0.0386	0.0456	0.0489	0.0500	0.0435	0.0431	0.0347	0.0692	0.0348	0.0069	0.0077	0.0091	**0.0050**	**0.0019**	0.0049	**0.0029**	**-0.0007**		0.0241
YOR	0.0883	0.0563	0.0361	0.0340	0.0351	0.0423	0.0450	0.0432	0.0417	0.0351	0.0656	0.0365	0.0054	0.0069	0.0086	**0.0012**	**0.0016**	**0.0057**	**0.0038**	**0.0030**	**0.0051**	
JAM	0.0932	0.0543	0.0318	0.0390	0.0438	0.0492	0.0472	0.0403	0.0430	0.0368	0.0708	0.0331	0.0051	0.0041	0.0081	**0.0048**	**-0.0008**	0.0037	**0.0056**	**-0.0005**	**0.0000**	**0.0026**
ROA	0.0925	0.0634	0.0436	0.0469	0.0518	0.0507	0.0551	0.0458	0.0546	0.0450	0.0753	0.0481	0.0125	0.0106	0.0157	**0.0020**	**0.0069**	0.0081	**0.0029**	**0.0010**	**0.0019**	**0.0040**
TAR	0.0865	0.0545	0.0363	0.0398	0.0440	0.0511	0.0518	0.0480	0.0472	0.0397	0.0671	0.0455	0.0090	0.0082	0.0134	**0.0039**	**0.0052**	0.0109	**0.0089**	**0.0048**	**0.0042**	**0.0056**
NEU	0.0864	0.0577	0.0384	0.0460	0.0511	0.0511	0.0557	0.0476	0.0487	0.0410	0.0669	0.0435	0.0101	0.0108	0.0132	**0.0063**	**0.0050**	0.0125	**0.0024**	**0.0028**	**0.0017**	**0.0052**
CF	0.0970	0.0651	0.0457	0.0475	0.0523	0.0554	0.0570	0.0533	0.0553	0.0492	0.0767	0.0497	0.0161	0.0144	0.0177	0.0084	**0.0137**	0.0160	0.0093	0.0067	0.0062	0.0085
WAC	0.0991	0.0686	0.0520	0.0566	0.0636	0.0666	0.0637	0.0595	0.0663	0.0571	0.0837	0.0561	0.0211	0.0202	0.0236	0.0135	**0.0107**	0.0169	0.0109	0.0081	0.0060	0.0137
COO	0.0937	0.0659	0.0485	0.0529	0.0566	0.0582	0.0624	0.0562	0.0633	0.0551	0.0804	0.0564	0.0180	0.0160	0.0207	**0.0078**	**0.0141**	0.0169	0.0084	0.0073	0.0080	0.0096
ED	0.0928	0.0630	0.0450	0.0528	0.0584	0.0602	0.0632	0.0556	0.0620	0.0543	0.0828	0.0547	0.0188	0.0172	0.0229	0.0143	**0.0140**	0.0151	0.0137	0.0092	0.0077	0.0143
SAV	0.1036	0.0730	0.0516	0.0536	0.0576	0.0547	0.0580	0.0564	0.0668	0.0563	0.0914	0.0546	0.0201	0.0179	0.0229	0.0094	**0.0155**	0.0206	0.0110	**0.0089**	0.0089	0.0086
ALT	0.0947	0.0661	0.0460	0.0515	0.0588	0.0648	0.0647	0.0608	0.0633	0.0555	0.0840	0.0519	0.0185	0.0178	0.0219	0.0091	**0.0086**	0.0185	0.0103	0.0072	0.0077	0.0126
STJ	0.1077	0.0795	0.0610	0.0655	0.0730	0.0781	0.0786	0.0729	0.0755	0.0692	0.0976	0.0697	0.0334	0.0316	0.0352	0.0212	0.0253	0.0297	0.0245	0.0199	0.0192	0.0253