2012

CHARACTERIZATION OF MICROSATELLITE LOCI AND PILOT POPULATION GENETIC ANALYSIS IN HICKORY SHAD, ALOSA MEDIOCRIS

FNU Vishakha
Virginia Commonwealth University

CHARACTERIZATION OF MICROSATELLITE LOCI AND PILOT POPULATION GENETIC ANALYSIS IN HICKORY SHAD, *ALOSA MEDIOCRIS*

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University

by

VISHAKHA

BACHELOR OF SCIENCE, SRM UNIVERSITY, INDIA 2010

MASTER OF SCIENCE, VIRGINIA COMMONWEALTH UNIVERSITY, 2012

Director: BONNIE L. BROWN, PH.D.
PROFESSOR, DEPARTMENT OF BIOLOGY

Committee Members:

Tracey Dawson Cruz, Ph.D., College of Humanities and Sciences
Derek M. Johnson, Ph.D., College of Humanities and Sciences
Karen M. Kester, Ph.D., College of Humanities and Sciences
Stephen P. McIninch, Ph.D., Center for Environmental Studies, Life Sciences
Aaron W. Aunins, Ph.D., VCU Department of Biology

Virginia Commonwealth University
Richmond, Virginia
August 2012

Acknowledgements

I am profoundly grateful to everyone who made this research possible. First and foremost, I thank my advisor Dr. Bonnie Brown for her endless patience, understanding, and encouragement. She has been a great mentor and confidante. I also thank Dr. Aaron Aunins for his guidance. He has always been there to answer any questions and has taught me many of the molecular techniques that I have used in this study. I thank my other committee members; Dr. Derek Johnson, Dr. Karen Kester, Dr. Tracey Dawson Cruz, and Dr. Stephen McIninch. They have been supportive of all my endeavors and I am truly honored to have them on my committee. I give a very special "thank you" to Dr. Jennifer Stewart for her optimism and invaluable support. Last but not the least, I thank my family, who have been incredibly supportive of my decision to study abroad and have been so very patient and loving.

Table of Contents

List of Figures

Page

List of Tables

Page

Abstract

CHARACTERIZATION OF MICROSATELLITE LOCI AND PILOT POPULATION GENETIC ANALYSIS IN HICKORY SHAD, *ALOSA MEDIOCRIS*

By Vishakha, B.S.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Biology at Virginia Commonwealth University.

Virginia Commonwealth University, 2012

Major Director: Bonnie L. Brown
Professor and Associate Department Chair, Department of Biology

The hickory shad (*Alosa mediocris*) is a relatively understudied species of the anadromous fish sub-family Alosinae. This study, the first population genetic analysis of this species, employed 12 neutral microsatellite loci to estimate genetic diversity and population structure in tributaries of lower Chesapeake Bay, Virginia including James River and its tributaries (Appomattox and Chickahominy Rivers), Rappahannock River, and Pamunkey River. Genetic variation was extremely low. Estimates of observed heterozygosity were lower than expected heterozygosity. Significant population structure was detected among the six samples (Φ_{ST} = 0.093, p = 0.01). Effective population sizes were low (N_e ranged from 2 to 134). The lack of genetic diversity, especially compared to that of the American shad, was striking and could be the result of a bottleneck that took place more than thirty years ago which may plausibly account for the low genetic variation observed across all populations.

Introduction

A Brief Introduction to the Hickory shad, Alosa mediocris

Clupeidae, a family of mostly marine foraging fish, support commercially valuable fisheries around the world. Although predominantly marine, 29 species of the Clupeidae are diadromous (McDowall 2003). First introduced by Myers in 1949, the term diadromy is used to refer to life history strategies of fish which migrate between marine and freshwater biomes at different stages in life (Myers 1949). Of the several diadromous clupeid species, many belong to the subfamily Alosinae, which includes the shads and river herrings. These species are anadromous, a special form of diadromy where the adults live in the ocean for a majority of their lives, but ascend freshwaters to spawn. These anadromous species play an important role in shaping freshwater ecosystems, as the adults deliver annual fluxes of marine derived nutrients to freshwater systems (Garman 1992).

Although shads are one of the most economically valuable fish in the world, on the east coast of the United States four species of *Alosinae,* the American shad *Alosa sapidissima,* the hickory shad *A, mediocris,* the alewife *A. pseudoharengus* and the blueback herring *A. aestivalis* have each been the mainstay of historically important fisheries and have generated considerable interest in the scientific community (McBride 2007). The least studied of the tetrad is the hickory shad which is intermediate in size between the American shad and the comparatively smaller alewife and blueback herring, the latter two also referred to as the river herrings (Mansueti

1962). Other distinct physical characteristics of the hickory shad are its large, superior mouth with a strongly projecting mandible which enters into the dorsal profile, and the relatively small number of gill rakers (19-21) observed on the lower limb of its first pharyngeal arch, whereas its sister species in general have 25 or more (Hildebrand and Schroeder 1928).

The hickory shad, *A. mediocris,* is widely distributed along the Atlantic coast of North America from Maryland to Florida (Harris *et al.* 2007). The taxonomic status of the hickory shad has been in a state of flux since the genus *Alosa* was split into three genera by Regan in 1917: *Alosa, Caspialosa* Berg, and *Pomolobus* Rafinesque. The hickory shad, along with its sister species the alewife and the blueback herring were classified under the genus *Pomolobus.* However, the works of Svetovidov (1964) and Bailey *et al.* (1954) combined *Pomolobus, Caspialosa* and *Alosa,* leading to a change in the scientific name of the hickory shad from *Pomolobus mediocris* to *Alosa mediocris* (Bowen *et al.* 2007). Interestingly, whereas its more celebrated sister species are experiencing grave declines despite conservation efforts along the entire east coast of North America (Waldman and Limburg 2009), the hickory shad appears to be increasing in abundance in various regions.

Historically, the hickory shad was considered to be the most cryptic of all the clupeids. It was speculated that hickory shad was a hybrid of *A. sapidissima* and one of the river herrings, *A. pseudoharengus* or *A. aestivalis* (Mansueti 1958, Mansueti and Kolb 1953). Because no information on the spawning history of hickory shad was available, some fish culturists attempted to cross these species in shad hatcheries to determine the viability of hybrids but without much success (Mansueti 1962). Moreover, ichthyologists debated over the spawning site of the hickory shad. Hildebrand and Schroeder (1928), based on their collections from many

parts of Chesapeake Bay over all seasons, concluded that the hickory shad does not advance to the freshwaters of the Chesapeake Bay to spawn. However, Mansueti (1962) established that the hickory shad ascends the freshwaters of Chesapeake Bay to spawn through successful hatching of eggs from an adult hickory shad female which were fertilized with a male from the freshwater portion of Chesapeake Bay. He also documented and described the developmental stages of the hickory shad in great detail, and provided morphological evidence to conclude that the Hickory shad is not a hybrid of its sister species (Mansueti 1962). However, there are no published genetic studies to date that corroborate the findings of Mansueti.

Although genetic diversity is usually accepted as one of the pivotal properties for the function of ecological communities and their resilience to alteration of the environment, some scientists also have attributed human-mediated hybridization as a key factor in determining patterns of biodiversity. In 2010, Coscia *et al.* (2010) corroborated that the two main species of shad widely distributed along the North Atlantic Ocean coasts from Morocco to Iceland (twaite shad, *A. fallax,* and allis shad, *A. alosa)*, are hybridizing as a result of river fragmentation through dam construction and habitat loss. Anthropogenic environmental perturbations such as dams disturb the patterns of fish migration and alter spawning behavior, i.e. fish from related species are forced to share spawning grounds thereby heightening the chances of inter-species hybridization (Coscia *et al.* 2010). Hence, the possibility that the hickory shad is indeed a hybrid of the American shad and one of the river herrings warrants genetic investigation.

Although the biology of the american shad *Alosa sapidissima* in different river systems has been widely studied (Limburg *et al*. 2003), very few studies have been conducted on the biology of the hickory shad in any river. Harris *et al*. (2007) studied the life history of hickory

shad in the St. Johns River, which is the longest river in Florida and is home to the southernmost population of the species. They documented several life history differences between the hickory shad population in St. Johns River in the early 2000s versus data from the 1970s. Based on other unpublished studies, Harris *et al*. (2007) hypothesized that there may be a latitudinal pattern in the timing of the spawning migration of Hickory shad, wherein the hickory shad populations in southern river systems tend to spawn earlier than those in more northerly water bodies. Most recently, Murauskas and Rulifson (2011) sampled hickory shad for two consecutive years from locations along the coasts of North Carolina to examine the reproductive development of hickory shad during its spawning migration. Their study indicated that overall, reproductive development in migrating hickory shad is contingent upon several factors such as size, age, energy reserves, geographic location, and time of year (Murauskas and Rulifson 2011). These life history traits should affect the population patterns exhibited by hickory shad, namely effective migration rate, gene flow, and population structuring among rivers.

Initially, Mansueti recorded elementary information related to the spawning of hickory shad such as the time of hatching post-fertilization. According to his study, hickory shad eggs are initially semi-agglutinant and semidemersal, lose their adhesive nature with age and can readily float in rapidly flowing water (Mansueti 1962). However, no information relating to the macro- and micro-habitat requirements for spawning in hickory shad was available, preventing the development of a habitat suitability model. Due to the increased interest in this relatively uncelebrated species in recent times, Harris and Hightower (2011) performed a study characterizing the spawning habitat of hickory shad and proposed a rudimentary habitat suitability model for the conservation and management of hickory shad, wherein the primary

parameters for habitat suitability were water velocity, temperature and substrate. Long term loss of suitable habitat could have population consequences such as inbreeding depression, reduced effective population size, and increased probability of extinction.

Although hickory shad populations have not been monitored adequately, it is anecdotally believed that hickory shad populations suffered a bottleneck in the late seventies (Atlantic States Marine Fisheries Commission (ASMFC) 2009). In 1980, the state of Maryland imposed a moratorium on the harvest of hickory shad from Maryland waters of the Chesapeake Bay due to severe declines. However, in recent years, the works of Batsavage and Rulifson (1998) indicate that the numbers of hickory shad are burgeoning in the Albermarle Sound, North Carolina, which has also lead to an increase in commercial fishing in the Albermarle Sound. Similarly, stock assessment reports of the ASMFC suggest that since the mid-1990s, hickory shad populations have experienced a surge in upper Chesapeake Bay and its tributaries (ASMFC 2009). These findings also are supported by landings data from the U.S. National Marine Fisheries Service (NMFS), which reported that there has been a ten-fold increase in the hickory shad landings from 5.6 metric tons in 1990 to 61.9 metric tons by 1999 (Waldman and Limburg 2003), though it is unclear if this increase in catch is a result of fishing effort shifting away from less abundant species such as American shad.

Population genetics is a branch of biology that allows estimation of the genetic composition of a species using molecular genetic characters such as allozymes, mitochondrial DNA, microsatellite loci, and single nucleotide polymorphisms (SNPs). A population genetic study can provide information on the gene richness, genetic diversity, migration, and phylogeny of a species. Microsatellites are defined as tandem repeats of short (1-6 bp) DNA motifs that are

present pervasively in eukaryotic genomes. These repetitive DNA sequences are generally located outside coding regions of DNA and are therefore thought to be selectively neutral. As a result of their abundance, selective neutrality, and high levels of polymorphism, microsatellites are widely used as markers in diverse fields such as association studies, population genetics, and forensics (Kelkar *et al*. 2010). Therefore, microsatellites are an ideal marker choice for generating information that enhances our knowledge of the genetics of hickory shad, setting the stage for a better understanding of its recent apparent success in the fragmented tributaries of the Chesapeake Bay system. As no genetic study on the hickory shad is heretofore available, the purpose of this study was to estimate and compare genetic variation among hickory shad populations of the different rivers of the Virginia portion of Chesapeake Bay by developing a suite of microsatellites from hickory shad and assessing cross-species amplification with American shad microsatellites. It was hypothesized, based on the limited life history and population data available, that hickory shad would exhibit population genetic parameters similar to patterns exhibited by American shad.

Materials and Methods

Sample collection

Hickory shad samples were collected by Virginia Department of Game and Inland Fisheries (VDGIF) and their collaborators in 2001 from the spawning grounds of various rivers in Virginia's portion of the Chesapeake Bay watershed including Rappahannock River, Pamunkey River and its tributary the South Anna, and James River and its tributaries Appomattox and Chickahominy (Figure 1). Samples from James River also were collected in 1998, permitting a limited temporal analysis.

DNA Extraction

DNA was extracted from 0.25 cm^3 samples of muscle or fin clip tissue preserved in 70% ethanol and stored at -80^oC following a standard DNA extraction protocol. Upon elution in 75 µl of 0.25X TE, a portion of the DNA was diluted 1:10 for PCR and the remainder was stored at -80^oC.

Isolation and Optimization of Microsatellite Loci

Two methods were used to generate novel microsatellite loci. First, a microsatellite-enriched library was prepared from a genomic DNA pool of five hickory shad specimens in accordance with the methods outlined by Glenn and Schable (2005). Briefly, genomic DNA was digested

into 300-1000 bp fragments with *Rsa*I and ligated to SuperSNX24 linkers. The linker ligation product was amplified with PCR, and the PCR products were denatured and hybridized to a cocktail of biotinylated SSR oligos [$(AAAT)_8$, $(AACT)_8$, $(AAGT)_8$, $(ACAT)_8$, and $(AGAT)_8$]. Recovered single strands enriched for repeats were repaired by PCR, TA-cloned with the TOPO™ TA cloning kit for sequencing (Invitrogen), and 96 recombinant colonies were chosen for sequencing with M13 primers. Amplicons of 500-1200 bp were Sanger sequenced in the forward direction. Sequences were visually proofread in Chromas Ver 2.01 (Technelysium, Inc.), and screened for di-, tri-, tetra-, and pentanucleotide repeat containing sequences in the program MSATCOMMANDER (Faircloth 2008). Of these sequences, those which contained microsatellites were used to develop primers in Primer3 (Rosen and Skaletsky 2000), and were screened on 24 individual *A. mediocris* for polymorphism using a MegaBACE fluorescent genotyper (Amersham Biosciences, Inc.) and fluorescently labeled primers (Boutin-Ganache 2001).

Microsatellite loci also were developed from whole genome sequence data derived from a 100 ng pooled DNA sample comprised of four hickory shad specimens, one each from Appomattox, Chickahominy, James, and Rappahannock Rivers. A 200 bp barcoded library (the hickory shad was prepared alongside an *A. aestivalis* library for analysis on the same chip) was prepared using the Ion Fragment Library Kit and Ion Xpress Barcode Adapters kit (Life Technologies) according to the manufacturer's protocol and size-selected using E-Gel® Size Select 2% Agarose (Invitrogen). Template preparation was carried out with the Ion PGM 200 Xpress Template Kit (Life Technologies) and emulsified Ion Sphere™ particles were collected by centrifugation in a SOLiD® emulsion collection tray (Life Technologies). Upon separation

from the oil layer, enriched washed ion sphere particles were prepared for sequencing using the Ion PGM 200 Sequencing Kit (Life Technologies) according to the manufacturer's protocol, which included resuspension in annealing buffer, sonication, and loading of approximately one-third of the material onto a 316 chip. The chip was analyzed using a Personal Genome Machine™ (PGM™) sequencer per the manufacturer's protocol. Barcoded individual sequence reads specific to *A. mediocris* were *de novo* assembled using Genomics Workbench Ver 5 (CLC Bio). The resulting contigs were sorted based on size, whereupon those larger than 200 bp were analyzed further using the program msatcommander (Faircloth 2008) to locate tandem repeats within sequences. After successful identification of microsatellite containing contigs, the contigs were further analyzed for suitable primer sites using Primer3 (Rosen and Skaletsky 2000). Candidate loci were tested to ensure amplification of the expected amplicon size using generic PCR conditions, fluorescent labeled primers (Boutin-Ganache *et al*. 2001), and annealing temperatures in accordance with the calculated primer melting temperatures. Optimized loci were screened for polymorphism using 24 individual *A. mediocris*, and a selection of those loci that produced at least three different alleles were utilized for genotyping.

Testing cross-species amplification of A. sapidissima microsatellite loci in A. mediocris

A set of eight hickory shad specimens (two each from Appomattox, Chickahominy, James, and Rappahannock Rivers) was amplified using nine primer pairs designed for microsatellite loci of the American shad. These nine loci included four developed by Waters et al. (2000)(Asa-4, Asa-6, Asa-8, Asa-9), and five loci (Table 1) developed by Julian and Bartron (2007) (AsaB020, AsaD029, AsaD031, AsaC249 and AsaD312). Using the published thermal cycling parameters,

the 5 loci that yielded polymorphic amplicons (Asa-4, Asa-9, AsaD029, AsaD031, and AsaC249) were fluorescently labeled and retained for genotyping.

PCR Amplification and Genotyping

PCR amplification of *A. mediocris* individuals was carried out in 6 µL reactions, each containing 3 µL GoTaq™ mastermix (Promega, Inc.), 0.6 µL primer mix (0.5 µM each), 1µL 1:10 DNA template, and 1.4 µL nuclease-free water. Amplicons were diluted 1X with distilled water and resolved using capillary electrophoresis in a MegaBACE 1000 fluorescent genotyper (Amersham Biosciences, Inc.). Up to three loci were pooled based on non-overlapping size classes in 96-well plates and each well included 0.50 µL of MapMarker 400 molecular size standard (Bioventures, Inc.), 8.5µL 0.1% Tween-20 solution, and not more than 0.5µL each of 1X diluted PCR products. Allele sizes were determined using Fragment Profiler (Amersham Biosciences, Inc.). Binsets were designed separately for each locus and were used for allele size calling for all samples, following which all samples were checked for scoring accuracy.

Population genetic analyses

Alleles for each locus were screened in the program Microchecker (Oosterhout *et al*. 2004) for evidence of null alleles or scoring errors. The program CONVERT (Glaubitz 2004) was used to produce the required data input formats for various population genetic analysis programs. WHICHLOCI, a program that ranks candidate microsatellite loci in descending order based on their discriminatory power, was used to evaluate the usefulness of the twelve microsatellite loci for population genetic analyses (Banks *et al.* 2003).

Tests for gametic disequilibrium as well as conformance to from Hardy-Weinberg Equilibrium (HWE) expectations were performed in Genepop (Raymond and Rousset 1995) using the default Markov –chain parameters. A Sequential Bonferroni correction was applied to the results from gametic disequilibrium. Fisher's exact tests were used to evaluate conformance to Hardy-Weinberg Equilibrium for each locus as well as over all loci for each population. A significance level of $\alpha = 0.05$ was used to appraise all statistical tests.. Observed heterozygosity, expected heterozygosity, observed heterozygosity, effective number of alleles, and the number of private alleles were calculated in GenAlex 6 (Peakall and Smouse 2006).

Whether there was evidence of recent effective population size reduction was assessed using BOTTLENECK (Version 1.2; Cornuet and Luikart 1997) under the two-phased model (T.P.M.) using the default settings. The BOTTLENECK program is based on the principle that the heterozygosity excess will be observed in populations which have suffered a recent bottleneck, i.e., observed heterozygosity will be larger than the expected heterozygosity if the microsatellite loci were in mutation-drift equilibrium. The BOTTLENECK program allows for microsatellite loci to evolve under one of three different mutation models specified by the user. The Infinite-Alleles Model (I.A.M.) is better suited for dinucleotide loci which takes into account multi-step changes in allele size whereas the second model, called the Stepwise Mutation Model (S.M.M.) is a much stricter one-step model suited for trinucleotide repeats or greater. The Two-step Mutation Model (T.P.M.), an intermediate between I.A.M. and S.M.M. Mode-shift analyses of allele frequency distribution and Wilcoxon sign-rank tests (Luikart and Cornuet 1997) also were performed to assess the possibility of recent population bottlenecks. Data for only the tri-,

tetra-, and pentanucleotide loci (10 of 12 loci) were analyzed in BOTTLENECK using the Wilcoxon test on the assumption that all loci fit the Two Phased model of mutation (T.P.M.).

Effective population size (N_e), defined as the effective number of breeding individuals in a population, and its associated confidence intervals were calculated for each population using the linkage disequilibrium method of Waples (2006) as implemented in LDN_e (Waples and Do 2007). F_{ST} estimates for all population pairs, which are indicative of the proportion of the total genetic variance contained in a population relative to the total genetic variance, were calculated on Genepop (Raymond and Rousset 1995). Once again GenAlex ver. 6 was used to estimate Φ_{ST} estimates based on AMOVA as an indication of overall population genetic structure (Peakall and Smouse 2006). Population structure among populations was also examined using a Bayesian approach in STUCTURE (Pritchard *et al*, 2000). In STRUCTURE, we used the correlated allele frequencies and admixture model and left other settings at the default values. STRUCTURE was ran for values of K (the number of clusters) from 100000 – 200000, with 5 iterations for each value of K.

Effective migration rate (N_em) for all population pairs was estimated using the private alleles method on GENEPOP (Slatkin 1985). Nei's genetic distance (D_s), which is based on the assumption that genetic divergence among populations arise due to genetic drift and mutation, was calculated for all population pairs using the online program population to population genetic distance calculator (Brzustowski 2012). Lastly, correlation between the geographic distance and the genetic distance matrices was tested using the Mantel test as implemented by GenAlex6 (Peakall and Smouse 2006).

Results

From the enriched cloned *A. mediocris* library, 80 clones were Sanger sequenced, of which 18 contained tandem repeats. Following testing of 13 loci, three produced at least three different alleles and were further analyzed and utilized for genotyping. The next-generation sequencing data resulted in 24,256 contigs ≥ 200 bp in length, of which 129 contained microsatellites. From those 129 contigs, 37 possessed sufficient flanking sequence for primer design and upon testing, yielded the expected amplicon size. Four microsatellite loci obtained from the next-generation sequences were optimized and utilized for population genetic analysis in this study.

A total of 311 hickory shad were collected from the tributaries of Chesapeake Bay, with 152 samples examined from the James River (n = 65 in 1998 and n = 87 in 2001), 42 samples examined from the Rappahannock River, 27 samples examined from the Appomattox River, 15 samples examined from the Pamunkey River, and 10 samples examined from the Chickahominy River wherein all samples were collected in 2001. However, of the 311 samples, 91 samples were missing data at more than three loci. Appomattox River 2001 population was missing data at locus Asa249 and James River 1998 population was missing data at locus Ame6882. Population genetic analyses were performed using both datasets (with and without missing data) and there were negligible differences in the results obtained from both datasets.

Hickory shad populations from the different rivers and tributaries of Chesapeake Bay

exhibited expected heterozygosity ranging from H_e = 0.33 (Pamunkey River 2001) to 0.40 (James River 2001), in every case higher than the corresponding observed heterozygosity which ranged from H_o = 0.29 (Appomattox River 2001) to 0.38 (James River 2001). The effective number of alleles ranged from A_e = 1.78 (Chickahominy River 2001) to 1.94 (Appomattox River 2001) (Table 4). Between James River 1998 and James River 2001 hickory shad populations, there was no remarkable change in genetic diversity as evidenced by these measures, although the allelic distributions for these two samples from the James River population were very different (described below). The number of private alleles ranged from 0.00 (Pamunkey 2001) to 2.17 (James 2001).

Only 4 of the 12 microsatellite loci were found by WHICHLOCI to contribute substantially (i.e., generated > 10% of the discriminatory power) to the population genetic analysis whereas the remaining 8 microsatellite loci contributed less than 10% each. The four most informative microsatellite loci (Ame15296, Asa4, Ame6882, and Asa31) also were the most polymorphic of those tested on hickory shad (Table 3). Of these, two were microsatellite loci designed for the American shad and the other two were derived from hickory shad whole genome sequences.

Of the six hickory shad populations, three populations (James River 1998, James River 2001, Rappahannock River 2001, and Appomattox River 2001) deviated significantly from the expectations of Hardy-Weinberg Equilibrium (Table 3). The remaining populations that conformed to HWE expectations were those represented by very small sample sizes. Linkage disequilibrium test results indicated that of 330 possible tests for linkage disequilibrium among pairs of loci within populations (5 populations x 65 pairwise comparisons = 330), none remained

significant after sequential Bonferroni correction. MICROCHECKER suggested the presence of null (non-amplifying) alleles at six of the twelve loci for one or two Hickory shad populations (Table 2). However, the indication of heterozygote deficits due to possibility of null alleles appeared to be arbitrary. Since no pattern could be deduced, the revelation of possible null alleles at these loci was not supported by errors in PCR or genotyping. Furthermore, there was no indication of large allele drop out, and thus all loci were retained for further analyses.

Estimates of population differentiation using Fisher's Exact Test indicated that the James River 1998 population of hickory shad was significantly different from all other populations ($X^2 = \infty$, $p < 0.001$, $df = 22$). The Φ_{ST} values and F_{ST} indices were comparable for most population comparisons, although Φ_{ST} values were generally larger than F_{ST}. Overall, AMOVA indicated that there was 9% molecular variance among populations and 91% within populations (Table 6). When the James River 1998 population was excluded from the analysis, a 7% molecular variance was indicated using AMOVA. Similarly, tests of pair-wise differentiation were relatively large among most population comparisons (Table 6). However, analysis of population structure based on a Bayesian approach using STRUCTURE (Pritchard *et al* 2000) failed to delineate population structuring among the hickory shad populations. Instead, substantial variation in the percent membership to lineages in individuals within populations was observed for values of K (the number of clusters) from 100000 – 200000, with 5 iterations for each value of K (Figure 2).

Estimates of effective population size (Table 5) were approximately in the same range (69 to 76) for four populations (Appomattox River 2001, James River 2001, Pamunkey River 2001, and James River 1998). The highest estimate of effective population size ($N_e = 135$) was

observed for the Rappahannock River 2001 hickory shad population. The lowest estimate (N_e = 2) was observed for the Chickahominy River 2001 hickory shad population; however the low sample size necessitates caution interpretation. Moreover, with the exception of James River 2001, all populations had upper confidence limits that included infinity suggesting low precision in N_e estimates.

Estimates of effective migration rate (N_em), calculated using private alleles method, for hickory shad population pairs were relatively high, approaching panmixis between some populations (Table 7) These estimates would have been biased, if F_{ST} values were used to calculate the N_em (Whitlock, 1999). Nei's genetic distance values, ($D_{a,}$ which may vary from 0 to 1) for all pairs of populations were relatively high (D_a = 0.12 for Appomattox 2001 and Pamunkey 2001), akin to the Φ_{ST} estimates. Mantel test for correlation of geographic and genetic distance matrices indicated that there was no correlation between the two measures (p = 0.14).

Prior to running BOTTLENECK to test for evidence of recent severe population reduction, two loci (Ame19 and Ame5315) were excluded from the analysis as they were dinucleotide repeats which better fit the Infinite Allele model (I.A.M.). Using the remaining ten loci, no evidence was found for a recent bottleneck in any of the six populations of Hickory shad that were tested. Chickahominy River 2001 hickory shad population indicated a shifted-mode which is only one of the three possible tests for recent bottlenecks. However, this was not corroborated by the Wilcoxon test.

Discussion

The current study on hickory shad populations employed next-generation sequencing in addition to conventional cloning and enrichment strategies to develop microsatellite loci for this species. Overall, the microsatellite panel consisted of 4 loci from next-generation sequencing, 3 loci from the conventional method, and 5 loci derived from *A. sapidissima* (Table 1). These are the first documented microsatellite loci designed for hickory shad. Also of note, from the original 37 loci that were found to be polymorphic, there are 12 additional loci that have not yet been completely optimized.

This pilot population genetic study on the hickory shad provides a sweeping idea of the genetic variation in several hickory shad populations of the Chesapeake Bay tributaries. Of the four most informative loci (Ame15296, Asa4, Ame6882, and Asa31), two were derived from American shad and two from hickory shad. The usefulness of cross-species loci is not surprising as utilization of microsatellites designed from a related species is a common practice in population genetics. In fact, cross-species amplification of microsatellites has also been used for clarification of "potential hybridizations" between related species. In 2011, DiBattista and Feldheim (2011) developed eight microsatellite loci for the ornate butterflyfish (*Chaetodon ornatissimus*) and successfully cross-amplified them in a sympatric sister species the scrawled butterflyfish (*Chaetodon meyeri*). Hybrids of these two species have been documented in eastern Indian Ocean populations but it is speculated that these two species might be hybridizing in other

regions of overlap. The eight polymorphic loci developed amplified reliably in both species and were shown to be useful to examine potential hybridization in other areas (DiBattista and Feldheim 2011). The reliable amplification of American shad microsatellites in hickory shad, alewife, and blueback herring might likewise be advantageous and this level variation exhibited by hickory shad is promising for a successful taxonomic study that might further elaborate on the evolutionary relationships among these four species.

It is interesting that a similar approach taken for alewife, in two different studies gave equivocal results. Kuhn and Kornfield (2004) investigated genetic diversity in two different populations (Albany and Newburgh) of alewife in the Hudson River, New York using American shad microsatellites. For the six loci typed, both Hudson River populations exhibited relatively high number of alleles (N_a = 8.8 to 10.3). The expected heterozygosity and observed heterozygosity estimates for the Albany population were H_e = 0.63 and H_o = 0.60. The expected and observed heterozygosity estimates for the Newburgh population were relatively higher than for Albany but also indicated conformance to HWE, at H_e = 0.75 and H_o = 0.74 (Kuhn and Kornfield 2004). Conversely, another study (Chilakamarri 2005) testing cross-species amplification of American shad microsatellites and examining the genetic heterogeneity in alewife sampled from two different sites in Connecticut (Bride Brook and Roaring Brook) and one site in Lake Michigan revealed low levels of genetic diversity. The mean observed heterozygosity for Bride Brook and Roaring Brook populations in Connecticut were H_o = 0.22 and H_o = 0.31, respectively, and even lower for the Lake Michigan alewife population H_o = 0.11. These are comparable to the low levels of genetic diversity observed in the current study of hickory shad populations from different rivers and tributaries of Chesapeake Bay. These trends

observed in alewife may have a similar derivation as the trends observed in hickory shad

In general, low levels of genetic heterogeneity may be attributed to high reproductive variance as a result of hatchery effects in populations that have been stocked or to genetic drift caused by a small effective population size (Waples *et al.* 1990, Christie *et al.* 2012). Although there have been efforts to restore hickory shad populations through hatchery supplementation in the Maryland portion of the Chesapeake Bay (Richardson *et al.* 2009), no restoration efforts have been documented for the hickory shad populations of Virginia's lower regions of the Chesapeake Bay, which were sampled for this study. Hence, the possibility of the low genetic variation in these populations as a result of hatchery effects can be dismissed. The lower limits of effective population size estimates were in general, low for all hickory shad populations examined in this study. Though the inclusion of infinity as the upper confidence interval limit for the N_e estimates for most populations reflects low accuracy, the relative numbers should be informative since the same parameters were used to estimate the effective population sizes for all populations. Hence, the current data indicate that low effective population size may account for the low genetic variation observed. The lack of genetic diversity, especially compared to that of the American shad, is striking and could also be the result of a putative bottleneck that took place more than thirty years ago. Leberg (2002) established that the mean number of alleles per locus for a species once greatly reduced after a severe bottleneck, never recovers even after numerous generations, along this does not take into account post-bottleneck migrations (Leberg 2002). Perhaps hickory shad went through more severe (or more frequent) bottleneck(s) than American shad and never recovered their lost allelic diversity.

In this study, the three populations with relatively large sample sizes (roughly $n \geq 40$ for

12 loci with an average of 6 alleles each) did not conform to the expectations of HWE. There are a number of factors that would cause populations to deviate from HWE. For example, despite the lack of reliable monitoring, it is anecdotally believed that hickory shad populations suffered a demographic bottleneck in the late-seventies (ASMFC, 2009). Such an event could have severely reduced the genetic diversity of hickory shad populations, making them increasingly genetically homogenous and could possibly have caused linkage disequilibrium. Wang et al. (1998) established that linkage disequilibrium among non-additive loci, occurring after a bottleneck can increase the genetic variance, thereby affecting the HWE. However, two factors indicate that this explanation should be discounted. First, the hickory shad populations in this study did not exhibit significant linkage disequilibrium. Second, roughly seven generations have transpired since the putative bottleneck, which is a sufficient time to restore HWE. Other factors that could be considered as possible causes of deviation from HWE, but for which there are no current data, include unequal numbers between the sexes (or unequal numbers males and females breeding), age structure effects (e.g., overlapping generations), reproductive variance (a.k.a. non-random family size), migration/emigration, and selection. Any of these could cause the observed deviation from HWE, result in high levels of variance, and also may affect subsequent statistical treatments such as AMOVA.

Despite the very low levels of heterozygosity observed in Chesapeake Bay hickory shad populations, significant differences among the populations of hickory shad were revealed by AMOVA (Φ_{ST} = 0.09), and Nei's genetic distance estimates. The Φ_{ST} values were generally similar to what was expected from the F_{ST} analysis. These levels were similar to differentiation observed among alewife in NY (F_{ST} = 0.04 Kuhn and Kornfield, 2004). But Waples (1998)

inferred that low effective population size estimates for populations biases their F_{ST} analysis, leading to larger F_{ST} estimates. However, the variation:variance relationship observed for hickory shad was the inverse of that exhibited by American shad populations in the Chesapeake Bay, which have high levels of genetic variation but only a 1% variance among populations (Aunins, 2010), whereas hickory shad exhibited low levels of genetic variation and quite high levels of differentiation. Interestingly, efforts to delineate population boundaries for hickory shad using STRUCTURE were unsuccessful; a result that is not unexpected given the high levels of effective migration indicated by private alleles estimates (approaching panmixis between some populations). A high rate of migration among populations (as observed here) reduces population structuring, increasing homogeneity across populations. Instead of population structuring, as was expected from the high Φ_{ST} values, substantial variation in lineages was observed within populations. The most reasonable explanation for the non-concurrence of Φ_{ST}, N_em, and STRUCTURE may simply be that too many assumptions of all tests were violated by the population data in the current study for the tests to be precise. These assumptions include adherence to the island model of migration (Wright 1943), constant N_e, random breeding within subpopulations, non-overlapping generations, and lack of selection and mutation effects. Several population samples were excessively small (n = 10 and n= 15 for the Chickahominy and Pamunkey River 2001 hickory shad populations, respectively) and this is reflective of small census size (S. McIninch, personal communication) and therefore effective population size. Inbreeding was likely present, albeit not severe, indicating some level of non-random mating. Of particular concern is the possibility that the rates of non-random mating or inbreeding might have differed among populations, and we have no data to evaluate these possibilities.

The lack of detectable among-population structure in hickory shad mirrors the finding of no structure in contemporary (year 2008-2009) populations of American shad of the Chesapeake Bay Rivers (Aunins, 2010). This loss of population differentiation is a relatively recent phenomenon, as subtle but significant population structure was observed for American shad in the James and Pamunkey Rivers in the early 1990s (Waters, 2000). This current trend differs from the findings for American shad in other North American regions. Hasselman *et al*. (2010) examined population structure of American shad from twelve Canadian water bodies using thirteen microsatellite loci. Using STRUCTURE, they inferred that the American shad sampled from the twelve rivers could be classified into seven clusters that were representative of genetically distinct groups. Significant differences in the in the life history strategies of American shad populations of Canada (Hasselman *et al* 2010) and the southern United States may have a role in the dissimilarity of patterns for Chesapeake Bay and Canadian populations of American shad. Whether such differences in population structuring over large spatial scales occurs for hickory shad remains to be seen as the current study is the only one that provides data relating to the extent of population differentiation in hickory shad.

According to the ASMFC Shad and River Herring Interstate Fisheries Management Plan (ASMFC 2009), due to difficulties in distinguishing hickory shad from American shad and river herrings, it has been difficult to gather reliable data on population trends of this sidelined species. In 1980, the state of Maryland imposed a moratorium on both American shad and hickory shad due to the difficulty in reliably differentiating these species in the field. Recently, landing reports of 11,000 pounds of hickory shad caught in Chesapeake Bay were deemed erroneous when it was discovered that the collectors had actually caught gizzard shad (Maryland Department of

Natural Resources, 2004). Instances like these establish the necessity of proper monitoring of hickory shad populations. Like the Chesapeake Bay landing report, it is possible that the James River 1998 population from the current study, which exhibited significantly different allele frequencies from remaining populations sampled just 3 years, or 1 generation prior to 2001, may in fact be a different species. Our James River 1998 samples were likely not gizzard shad because our samples successfully amplified with American shad loci whereas Julian and Bartron (2007) reported that these loci did not amplify reliably from gizzard shad (Julian and Bartron 2007). One way to address whether our James River 1998 samples are from alewife or blueback is to complete additional analyses of known alewife and blueback populations each using these same 12 loci.

This study was not designed to examine whether the hickory shad is a hybrid; however, it would be interesting to examine the relationships among these four species concurrent with more detailed study of hickory shad life history. Recently, Coscia *et al.* (2010) utilized mitochondrial DNA and an array of microsatellite loci to appraise the prevalence of hybridization in populations of Allis shad and Twaite shad in four rivers of north-western Europe (Coscia *et al* 2010). They found that the two species, which have slightly different requirements for spawning conditions, are increasingly hybridizing as a result of dam construction and consequent lack of access to spawning grounds. These factors are in play in Chesapeake Bay tributaries as well and could have similar consequences for local alosine species. By more closely examining the taxonomic and ecological relationships among these species, we could better devise adaptive management strategies that facilitate spawning success of all four species. Developing additional microsatellite loci for the sister species is possible as we have next-generation

sequencing data from the other three alosine species inhabiting Chesapeake Bay. If hickory shad were or are currently affected by hybridization, the most promising approach for future work will be to use a combination of the maternally inherited mitochondrial DNA and 18S rRNA genes in conjunction with additional neutral microsatellite loci to decipher the enigma of the hickory shad's origin.

Being the first report on hickory shad population genetics, this work sheds light on the current status of this understudied species in the rivers and tributaries of Chesapeake Bay and raises questions about the population demographics that require further investigation. It is difficult to genetically characterize marine species adequately since in general, they violate many assumptions of population genetic testing. Hence, resampling from time to time is necessary in order to derive a complete picture of the status of such species with respect to population genetics (Waples 1998). Given the relatively low levels of allelic variation detected, the microsatellite loci used here indicate that a severe bottleneck or series of bottlenecks must have been experienced in the past. These low levels of genetic variation do not bode well for the long-term persistence of the hickory shad provided the observed levels of microsatellite (selectively neutral) diversity are indicative of the levels of adaptive genetic variability in these populations. Bekessy *et al* (2003) established that there is a poor correlation between neutral marker variation and the adaptive variation which is required for the fitness of a species (Bekessy *et al* 2003). Strand *et al.* (2011) used both neutral and major histocompatibility complex markers to asses the differences in the levels of genetic variation contained in small and fragmented grouse populations as opposed to larger ones (Strand *et al* 2011). Given that microsatellite markers give an inadequate picture of the evolutionary potential of a species, it

would be useful to test one or more markers for genes that are subject to selective pressure in conjunction with neutral DNA markers on the hickory shad populations to accurately assess their fitness.

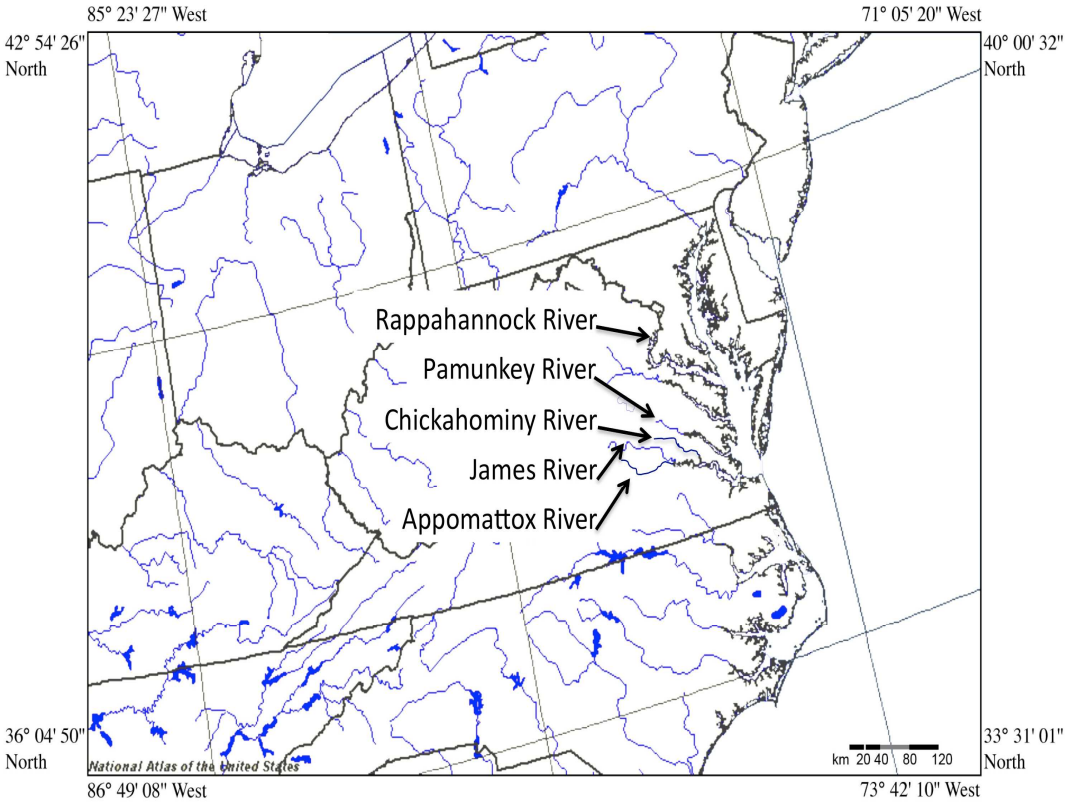

Figure 1: Map of the rivers and tributaries of lower Chesapeake Bay. Pamunkey River is a tributary of the York River (not shown here) and Appomattox River and Chickahominy River are tributaries to the James River.

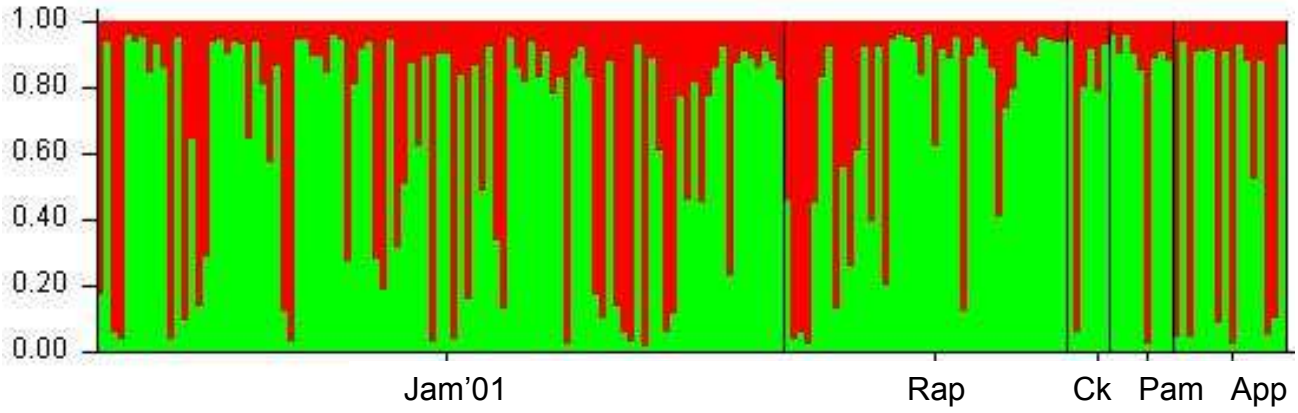

Figure 2: Barplot from K=2 clusters from the program STRUCTURE (Pritchard *et al.* 2000) for hickory shad samples collected from rivers and tributaries of the Chesapeake Bay. Each individual shad is represented by a single vertical bar and the percent association to each cluster is depicted by two different colors. Each population is separated by a vertical black line.

Table 1: Characteristics of the twelve microsatellite loci assayed on *Alosa mediocris*.

Locus Genbank Acc. No.	Repeat motif	Primer sequence	Size range (bp)
Asa4 AF039658	(ACC)2(AAC)12(AGC)6	F: TET-GAAGACAATACAGTAATAAACC R: GCGGGAGGCCAGACATA	110-180
Asa9 AF039661	(TTTC)7	F: FAM-GGGAATAAGGGATGTAGCCAAGAT R: AGGAGAAGGAAAGGGGAGTGAGAG	150-230
AsaD029 EF014997	(CTAT)20	F: HEX-ATTATGCACAGGAATCTGGAAG R: TGTGCTTACAAAAGTGACATGG	182–254
AsaD031 EF014999	(CTAT)14	F: HEX-TTCCTGATATTTCTTGTGAGGG R: ATTTCTGTGGAAACCTTTTGG	180-240
AsaC249 EF014994	(CATA)8(TTCT)13	F: FAM-TTATTACAACGGTGAATTGAGTG R: TAAGTGCATGTTGTGTGTGATG	243-367
Ame15296 TBD	(AGAGC)5	F: CCTGAGCGGATGGTGTAATC R: CAGTCGGGCGTCATCAGCAACTCTTCCGTCCAGC I: TET-CAGTCGGGCGTCATCA	150-180
Ame1808 TBD	(ACTTT)4	F: CAGTCGGGCGTCATCATGCAGTGATCGTGAAGCC R: TGGGCCACACACCTTTAGC I: TET-CAGTCGGGCGTCATCA	161-192
Ame5315 TBD	(GT)10	F: GGTGCGCTTCCTCTACAGC R: CAGTCGGGCGTCATCATGTACAAG GCCAGTCACCC I: HEX-CAGTCGGGCGTCATCA	149-181
Ame6882 TBD	(CCT)6	F: AATGATGTCGTATAATTCCAGGC R: CAGTCGGGCGTCATCAGGATGATTGTCAGTACTCCACC I: FAM-CAGTCGGGCGTCATCA	171-212
Ame64 TBD	(AGGT)4(AGAT)7	F: ATGTGCACCTGGGCAAGC R: CAGTCGGGCGTCATCACCTAGTCAGTCTTGAATTTCCTC I: FAM-CAGTCGGGCGTCATCA	155-180
Ame63 TBD	(AGTT)13	F:CAGTCGGGCGTCATCATCCAGCCTCACAACAGTCC R: CAAGGGCAAAGGCTTCCAG I: TET-CAGTCGGGCGTCATCA	270-295
Ame19 TBD	(GT)6	F:CAGTCGGGCGTCATCATTTCGGATGTGCAGAGGTTATAC R: GGTGAAACGGAGAACAGGC I: HEX-CAGTCGGGCGTCATCA	165-200

Table 2: Results of MICROCHECKER (Oosterhout *et* al. 2004) analysis of microsatellite genotypes collected from hickory shad in five rivers and tributaries of Chesapeake Bay. Yes and no stands for the presence and absence of null alleles respectively. NA indicates lack of data.

	Asa 9	Asa 4	Asa 31	Asa 29	Asa 249	Ame 64	Ame 63	Ame 19	Ame 5315	Ame 1808	Ame 15296	Ame 6882
Population												
Jam '98	no	no	no	yes	No	no	yes	no	no	no	yes	no
Jam '01	no	no	no	no	No	no	no	no	no	no	yes	no
Rap '01	yes	no	no	no	No	no	no	no	no	no	no	no
App '01	no	no	no	no	NA	no	no	no	no	yes	no	no
Pam '01	no	no	no	no	No	no	no	no	no	yes	no	no
Chk '01	yes	no	no	no	No	no	no	no	NA	NA	NA	NA

Table 3: WHICHLOCI (Banks et al. 2003) ranking of the twelve microsatellite loci used to examine *Alosa mediocris* in Chesapeake Bay tributaries.

Rank	Locus	Score	% (Relative Score)
1	Ame15296	0.4142	18.5808
2	Asa4	0.3077	13.8029
3	Ame6882	0.2959	13.272
4	Asa31	0.2781	12.4757
5	Asa9	0.1834	8.2287
6	Ame5315	0.1775	7.9632
7	Asa29	0.1479	6.636
8	Ame1808	0.142	6.3706
9	Ame64	0.1006	4.5125
10	Ame63	0.071	3.1853
11	Ame19	0.0651	2.9198
12	Asa249	0.0458	2.0524

Table 4: Summary data and population genetics analyses for hickory shad samples collected from major Chesapeake Bay tributaries. Calculated in GenAlex6 (Peakall and Smouse 2006): N = sample size, N_a= number of alleles, A_e= effective number of alleles, H_o= observed heterozygosity, H_e= expected heterozygosity. HWE analyzed using GENEPOP (Raymond and Rousset 1995).

Locus		App'01	Chk'01	Jam'01	Pam'01	Rap'01	Jam'98
Asa9	N	25	10	104	12	41	57
	N_a	5	5	8	4	6	7
	A_e	1.60	1.92	1.90	1.68	1.76	1.91
	H_o	0.36	0.50	0.48	0.42	0.32	0.49
	H_e	0.37	0.48	0.47	0.41	0.43	0.48
	HWE	0.12	0.27	0.10	0.55	0.01	0.39
Asa4	N	25	10	119	13	39	41
	N_a	6	3	12	5	8	6
	A_e	4.27	2.94	4.09	3.71	4.31	4.19
	H_o	0.84	0.30	0.69	0.62	0.69	0.76
	H_e	0.77	0.66	0.76	0.73	0.77	0.76
	HWE	0.57	0.20	0.00	0.22	0.22	0.08
Asa31	N	21	9	96	12	36	56
	N_a	4	4	6	4	6	4
	A_e	3.65	2.190	2.99	3.10	3.21	3.21
	H_o	0.48	0.44	0.68	0.75	0.78	0.77
	H_e	0.73	0.54	0.67	0.68	0.690	0.690
	HWE	0.17	1.00	0.14	0.90	0.01	0.03

Asa29	N	18	10	120	13	41	58
	N_a	1	2	8	1	2	6
	A_e	1.00	1.11	1.15	1.00	1.03	1.22
	H_o	0.00	0.10	0.08	0.00	0.02	0.12
	H_e	0.00	0.10	0.13	0.00	0.02	0.18
	HWE			0.02			0.01
Asa249	N	0	8	118	11	41	57
	N_a	0	2	5	2	3	3
	A_e	0.00	1.13	1.39	1.20	1.48	1.26
	H_o	0.00	0.13	0.26	0.18	0.34	0.23
	H_e	0.00	0.12	0.28	0.17	0.33	0.21
	HWE			0.04		1.00	1.00
Ame64	N	19	7	124	12	40	49
	N_a	3	3	4	3	3	2
	A_e	1.24	2.09	1.37	1.19	1.32	1.51
	H_o	0.11	0.43	0.26	0.17	0.18	0.39
	H_e	0.19	0.52	0.27	0.16	0.24	0.34
	HWE		1.00	0.02	1.00	0.08	1.00
Ame63	N	16	7	118	11	37	42
	N_a	3	2	6	1	3	4
	A_e	1.14	1.15	1.18	1.00	1.09	1.28
	H_o	0.13	0.14	0.14	0.00	0.08	0.10
	H_e	0.12	0.13	0.15	0.00	0.08	0.22
	HWE	1.00		0.07		1.00	0.00
Ame19	N	18	6	113	9	36	45
	N_a	4	2	5	2	2	4
	A_e	1.87	1.39	1.63	1.70	1.35	1.69
	H_o	0.28	0.33	0.37	0.33	0.25	0.31
	H_e	0.47	0.28	0.39	0.40	0.26	0.41
	HWE	0.52		0.00	1.00	1.00	0.00

Ame5315	N	20	7	112	12	37	55
	N_a	5	2	7	3	4	8
	A_e	2.69	1.96	2.21	2.27	1.90	2.32
	H_o	0.70	0.57	0.66	0.42	0.46	0.67
	H_e	0.63	0.49	0.55	0.56	0.47	0.57
	HWE	0.00	1.00	0.00	0.11	0.00	0.00
Ame1808	N	22	8	121	14	41	62
	N_a	6	2	6	3	4	7
	A_e	3.21	1.97	2.22	2.18	2.13	2.16
	H_o	0.46	0.63	0.60	0.43	0.46	0.50
	H_e	0.69	0.49	0.55	0.54	0.53	0.54
	HWE	0.00	0.39	0.00	0.02	0.04	0.00
Ame15296	N	1	3	9	2	1	5
	N_a	1.00	2.00	1.20	1.10	1.00	1.32
	A_e	0.00	0.00	0.05	0.09	0.00	0.04
	H_o	0.00	0.50	0.17	0.09	0.00	0.24
	H_e		0.11	0.00			0.00
Ame6882	N	17	5	105	9	39	0
	N_a	2	2	7	2	2	0
	A_e	1.64	1.47	1.80	1.25	1.49	0.00
	H_o	0.18	0.00	0.34	0.22	0.26	0.00
	H_e	0.39	0.32	0.45	0.20	0.33	0.00
	HWE	0.37	0.11	0.00	1.00	0.13	

Table 5: Estimates of effective population size (N_e) of Hickory shad from different rivers of Chesapeake Bay using the program LDN_e (Waples and Do 2008).

Sample	N_e	95% CI
Appomattox 2001	76	(3, ∞)
Chickahominy 2001	2	(1, ∞)
James 2001	76	(29,19042)
Rappahannock 2001	135	(24, ∞)
Pamunkey 2001	73	(5, ∞)
James 1998	69	(26, ∞)

Table 6: Pairwise matrix of Φ_{ST} values (below diagonal) and F_{ST} values (above diagonal) for populations of hickory shad in lower Chesapeake Bay Rivers. Φ_{ST} values were generated in GenAlex6 (Peakall and Smouse 2006) within an AMOVA framework. F_{ST} values were generated in Genepop (Raymond and Rousset, 1995).

	App'01	Chk'01	Jam'01	Pam'01	Rap'01	Jam'98
App'01	--	0.11	0.08	0.10	0.09	0.15
Chk'01	0.11	--	0.02	0.03	0.03	0.08
Jam'01	0.10	0.04	--	0.01	0.01	0.07
Pam'01	0.07	0.00	0.00	--	0.01	0.09
Rap'01	0.23	0.13	0.07	0.05	--	0.08
Jam'98	0.19	0.06	0.09	0.06	0.20	--

Table 7: Pairwise matrix of effective migration rate (N_em, below diagonal) and Nei's genetic distance (D_a, above diagonal) for populations of hickory shad in lower Chesapeake Bay Rivers. N_em estimates were generated in Genepop (Raymond and Rousset 1995). D_a values were generated using population to population genetic distance calculator (Brzustowski 2012).

		D_s					
		App'01	**Chk'01**	**Jam'01**	**Jam'98**	**Pam'01**	**Rap'01**
N_em	**App'01**	--	0.17	0.12	0.21	0.12	0.12
	Chk'01	1.25	--	0.07	0.15	0.06	0.07
	Jam'01	5.69	2.86	--	0.11	0.04	0.03
	Jam'98	3.90	1.86	9.21	--	0.13	0.12
	Pam'01	4.63	1.23	2.86	4.59	--	0.03
	Rap'01	3.26	1.46	7.99	4.51	3.57	--

References

Atlantic States Marine Fisheries Commission (ASMFC). 2009. Amendment 2 to the Interstate Fishery Management Plan For shad and river herring. Annapolis, Maryland. 166 pages.

Aunins, A.W. 2010. Genetic evaluation of American shad, *Alosa sapidissima,* restoration success in James River, Virginia. Ph.D. dissertation. Virginia Commonwealth University.

Bailey, R.M., H.E. Winn, and C.L. Smith. 1954. Fishes from the Escambia River, Alabama and Florida, with ecologic and taxonomic notes. *Proceedings of the academy of natural science of Philadelphia*. **56**: 109-164.

Barton, N.H., and M. Slatkin. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. *Heredity*. **56**: 409-415.

Banks, M.A., W. Eichert, and J.B. Olsen. 2003. Which genetic loci have greater population assignment power? *Bioinformatics.* **19**: 1436-1438.

Batsavage, C.F., and R.A. Rulifson. 1998. Life history aspects of the hickory shad (*Alosa mediocris*) in the Albemarle Sound/Roanoke River watershed, North Carolina. ICMR Contribution Series No. ICMR-98-02.

Bekessey, S.A., R.A. Ennos, M.A. Burgman, A.C. Newton, and P.K. Ades. 2003. Neutral DNA markers fail to detect genetic divergence in an ecologically important trait. *Biological Conservation.* **110**: 267-275.

Boutin-Ganache, I., M. Raposo, M. Raymond, and C. Deschepper. 2001. M-13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. *Biotechniques.* **31**: 24-26.

Bowen, B.R., B.R. Kreiser, P.F. Mickle, J.F. Schaefer and S.B. Adams. 2008. Phylogenetic relationships among North American *Alosa* species (Clupeidae). *Journal of Fish Biology.* **72**: 1188-1201.

Brzustowski, J. 2012. Population to population genetic distance calculator. URL http://www2.biology.ualberta.ca/jbrzusto/GeneDist.php.

Chilakamarri, S.R. 2005. Genetic differentiation in alewife populations using microsatellite loci. M.S. thesis. Worcester Polytechnic Institute.

Cornuet, J.M., and G. Luikart, 1997. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics.* **144**: 2001-2014.

Coscia, I., V. Rountree, J.J. King, W.K. Roche, and S. Mariani, 2010. A highly permeable species boundary between two anadromous fishes. *Journal of Fish Biology*. **77**: 1137-1149.

Christie M.R., M.L. Marine, R.A. French, R.S. Waples, and M.S. Blouin. 2012. Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. *Heredity.* 00: 1-7.

DiBattista J.D., and K.A. Feldheim. 2011. Isolation and characterization of eight microsatellite loci in *Chaetodon ornatissimus* and cross-amplification in a sympatric sister species, *Chaetodon meyeri*. *Conservation Genetics Resources.* **3**: 213-216.

Faircloth, B.C. 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* **8**: 92-94.

Garman, G.C. 1992. Fate and potential significance of postspawning anadromous fish carcasses in an Atlantic coastal river. *Transactions of the American Fisheries Society.* **121**: 360-394.

Glaubitz, J.C. 2004. CONVERT: a user friendly program to reformat diploid genotypic data for commonly used population genetic software packages. *Molecular Ecology Notes.* **4**: 309-310.

Glenn, T.C. and N.A. Schable. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology.* **395**: 202-222.

Guo, S.W., and E.A. Thompson.1992. Performing the exact test of Hardy-Weinberg proportions for multiple alleles. *Biometrics*.**48**: 361-372.

Harris, J.E., R.S. McBride, and R.O. Williams. 2007. Life history of hickory shad in the St. Johns River, Florida. *Transactions of the American Fisheries Society*. **136**: 1463-1471.

Harris, J.E., and J.E. Hightower. 2011. Spawning habitat selection of hickory shad. *North American Journal of Fisheries Management*. **31**: 495-505.

Hasselman D.J., R.J. 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.

Hildebrand, S.F. and W.C. Schroeder. 1928. Fishes of Chesapeake Bay. *Bulletin of the Bureau of Fisheries.* **53**: 83-85.

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

Kelkar, Y.D., K.D. Makova, N. Strubczewski, S. Hile, F. Chiaromonte, and K.A. Eckert. 2010. What Is a Microsatellite: A computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC Repeats. *Genome Biology and Evolution*.2: 620–635

Kuhn, K. and I. Kornfield. 2004. Genetic differentiation of the alewife, *Alosa pseudoharengus*, in the Hudson River. Section V:20 pp. *In* W.C Neider and J.R. Waldman, (eds), Final Reports of the Tibor T. Polgar Fellowship Program, 2003. Hudson River Foundation.

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

Li, L., H.P.Wang, C.B.Givens, S. Czesny, and B.L. Brown. 2007. Isolation and characterization of microsatellites in yellow perch (*Perca flavescens*).*Molecular Ecology Notes*.**7**: 600-603.

Limburg, K. E., K.A. Hattala, and A. Kahnle. 2003. American shad in its native range. In *Biodiversity, status, and conservation of the world's shads* (Limburg, K. E. and Waldman, J. R. eds), pp. 125-140.Bethesda, MD: American Fisheries Society.

Limburg, K.E., and J.R. Waldman. 2009. Dramatic declines in North Atlantic diadromous fishes. *Bioscience*. **59**: 955-965.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. Genetics.**152**: 1753-1766.

Mansueti, R.J. and H. Kolb.1953. A historical review of the shad fisheries of North America. Maryland Department of Research & Education Chesapeake Biol. Lab. Publ. (97).

Mansueti, R.J. 1958. The hickory shad unmasked. *Nature*. **51**: 3651-3654

Mansueti, R.J. 1962. Eggs, larvae, and young of the hickory shad, *Alosa mediocris*, with comments on its ecology in the estuary. *Chesapeake Science* **3**: 173-205.

Maryland Department of Natural Resources (DNR) Fisheries Management Plan (FMP Alosines 2004) DNR, Annapolis, Maryland. Available: http://www.dnr.state.md.us/fisheries/management/FMP/FMPAlosines04.pdf

McBride, R.S., N.A. Trippel, M.S. Allen. (2007). Seasonal trends in abundance and size of juvenile American shad, hickory shad, and blueback herring in St. John's River, Florida, and comparison with historical data. *Transactions of American Fisheries Society.* **136**: 988-993.

McDowall, R. M. (2003). Shads and diadromy: implications for ecology, evolution, and biogeography. In *Biodiversity, Strategies, and Conservation of the World's Shads*. (Limburg, K. E. and Waldman, J. R.), pp. 11–23. Bethesda, MD: American Fisheries Society.

Murauskas, J.G. and R.A. Rulifson. 2011. Reproductive Development and related observations during the spawning migration of hickory shad. *Transactions of the American Fisheries Society.* **140**: 1035-1048.

Myers, G. S. 1949. Usage of anadromous, catadromous and allied terms for migratory fishes. *Copies*, **1949**: 89–97.

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

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

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

Raymond, M. and F. Rousset. 1995. An exact test for population differentiation. *Evolution.* **49**: 1280-1283.

Richardson, B.M., C.P. Stence, M.W. Baldwin, and C.P. Mason. 2009. Hickory shad restoration in three Maryland rivers. F-57 Segment 9 Progress Report, Maryland DNR. 48 pages.

Rozen, S., and H J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Humana Press, Totowa, NJ, pp 365-386

Slatkin, M. 1985. Rare alleles as indicators of gene flow.*Evolution.* **39**: 53-65.

Strand, T.M., G. Segelbacher, M. Quintela, L. Xiao, T. Axelsson, and J. Hoglund. 2011. Can balancing selection on MHC loci counteract genetic drift in small, fragmented populations of black grouse? *Ecology and Evolution.*

Svetovidov, A.N. 1964. Systematics of North American anadromous Clupeoid fishes of the genera *Alosa, Caspialosa* and *Pomolobus. Copeia.* **1964**: 118-130.

Waldman, J.R., and K.E. Limburg. 2003. The world's shads: A summary of their status, conservation, and research needs. Pages 363-369 *in* K.E. Limburg, and J.R. Waldman, editors. Biodiversity, status, and conservation of the world's shads. American Fisheries Society Symposium 35, Bethesda, MD

Wang, J., A.Caballero, and W.G. Hill. 1998. The effect of linkage disequilibrium and deviation from Hardy-Weinberg proportions on the changes in genetic variance with bottlenecking. *Heredity*. **81**: 174-186.

Waples, R.S. 1998. Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *The American Genetic Association*. **89**: 438-450.

Waples, R.S., G.A. Winans, F.M. Utter, and C. Mahnken. 1990. Genetic monitoring of Pacific Salmon hatcheries. NOAA Technical Report. 33-37

Waples, R.S., and C. Do. 2007. LDNE: A program for estimating effective population

size from data on linkage disequilibrium. *Molecular Ecology Resources.*

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

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

Whitlock, M.C., and D.E. McCauley. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Heredity.* 117-125.

Wright, S. 1943. Isolation by distance. Genetics **28**: 114-138.

Vita

Vishakha was born in Jamshedpur, India on 23 November 1987. She graduated from Delhi Public School, Vasant Kunj, New Delhi in 2005. She received her Bachelor of Science in Genetic Engineering from SRM University, India in 2010. Vishakha volunteered in a Cytogenetics lab before joining the Master of Science in Biology program at Virginia Commonwealth University, Richmond, Virginia. After graduating, she will teach Biology and intends to pursue further graduate study in life sciences.