GENETIC DIFFERENTIATION OF THE ALEWIFE, *ALOSA PSEUDOHARENGUS*, IN THE HUDSON RIVER

A Final Report of the Tibor T. Polgar Fellowship Program

Kristen L. Kuhn

Polgar Fellow

Marine Biology Program
School of Marine Sciences
University of Maine
Orono, Maine 04469

Project Advisor:

Irv Kornfield
School of Marine Sciences
University of Maine
Orono, Maine 04468

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.

ABSTRACT

Alewives (*Alosa pseudoharengus*) are native to the eastern seaboard of the United States and Canada. Alewife life histories include both anadromous and freshwater forms. Recent introductions of alewives into non-native, freshwater habitats have caused much concern. In the United States, introduced freshwater alewife populations have become established in at least eighteen different states. Approximately 42% of threatened and endangered species in the United States are at risk from non-indigenous species, and during the past century, exotic species have been a factor in 68% of such extinctions in the United States. This project focused on genetic differentiation of alewives (*Alosa pseudoharengus*) within the Hudson River. The examination of alewives from two locations in the Hudson River, as well as the examination of alewives from other rivers and freshwater habitats, permits estimation of genetic differentiation, both within the River and among other populations. I used six microsatellite markers that have been previously characterized for the congener, the American shad (*Alosa sapidissima*). The markers were employed to characterize and compare samples of alewives from critical locations, including anadromous populations from Maine and Canada, and the Finger Lakes of New York. The level of genetic differentiation between population samples in the Hudson River is relatively low when compared to that of other anadromous populations. When combined with other analyses, these data provide an overall idea of the genetic structure of alewives in the northeastern United States and Canada.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The alewife (*Alosa pseudoharengus*) is an anadromous clupeid species of eastern North America, with a range from Newfoundland, Canada to the Carolinas, USA. The anadromous alewife, also termed river herring, spends the majority of its adult life at sea, returning to freshwaters only to spawn. Spawning runs occur in a south to north progression, and begin from March to May (Neves 1981) as adults migrate upstream and spawn at water temperatures between 12°C and 16°C. Adults return to their natal rivers much like salmon and shad. The young generally remain in freshwater for two to four months until they reach a size of two to four inches in length, when they begin migrations to sea.

In addition to anadromous populations, numerous populations exist solely in freshwater. Freshwater populations occur in the Great Lakes and the Finger Lakes of New York, in addition to lakes in which alewives have been introduced. There is no compelling evidence that alewives resided in freshwater prior to the Pleistocene epoch. Freshwater populations exhibit similar life history strategies which were possibly acquired during colonization and isolation after glacial retreat.

Alewives were first reported in Cayuga and Seneca Lakes, and in the Erie Canal, in June of 1868 (Bean 1884, Smith 1890). Alewives were not observed in Lake Ontario until 1873 (Wright 1892). The section of the Erie Canal linking the Hudson-Mohawk Rivers to Lake Ontario was opened to barge traffic in 1825, thereby potentially allowing alewives from the Atlantic to access the Finger Lakes and Lake Ontario (Smith 1995). Atlantic salmon was the major predator in Lake Ontario and the St. Lawrence River until the 1860s, when its numbers sharply declined (Smith 1890). Smith (1995) linked this

decline to the entrance of alewives into Lake Ontario via the Erie Canal. Extirpation of Atlantic salmon may have occurred through the consumption of thiaminase-containing alewives (Ketola et al. 2000). The detrimental effects of exotic alewives on native populations in many freshwater environments, including the Great Lakes and Finger Lakes, have been well documented, especially with regard to thiaminase. Thiaminase is an enzyme that catalyzes the destruction of thiamine, resulting in anorexia, lethargy, nonresponsiveness to disturbance, exhaustion, deformed yolk sacs, congestion in the heart region, and hydrocephalus (Ketola et al. 2000). Abnormally low concentrations of thiamine in Atlantic salmon eggs are reportedly associated with a forage base that includes alewives, which are known to contain the enzyme thiaminase (Neilands 1947, Gnaedinger and Krzeczkowski 1966).

There are two possible marine routes of alewife invasion into the Great Lakes: the Hudson River (via the Erie Canal), and the St. Lawrence River. When the Erie Canal was complete, it was 585 miles long and consisted of 83 locks. Water movement in the Canal was episodic and highly variable, thus making it a string of 82 narrow, shallow ponds of equal width and depth, but varying in length (Daniels 2001). Via allozyme analysis, Ihssen et al. (1992) concluded that alewives in the Great Lakes were probably recent invaders from the Hudson-Mohawk Rivers via the Erie Canal and Finger Lakes. Ihssen et al. (1992) did not study alewives from the Hudson River and they did not present alternative hypotheses based on their data. The second route of proposed invasion into the Great Lake and Finger Lakes is from Maritime Canada via the St. Lawrence River. This hypothesis states that alewives are presumed to have invaded Lake Ontario via the Welland Canal. However, Bean (1892) reported that alewives were not

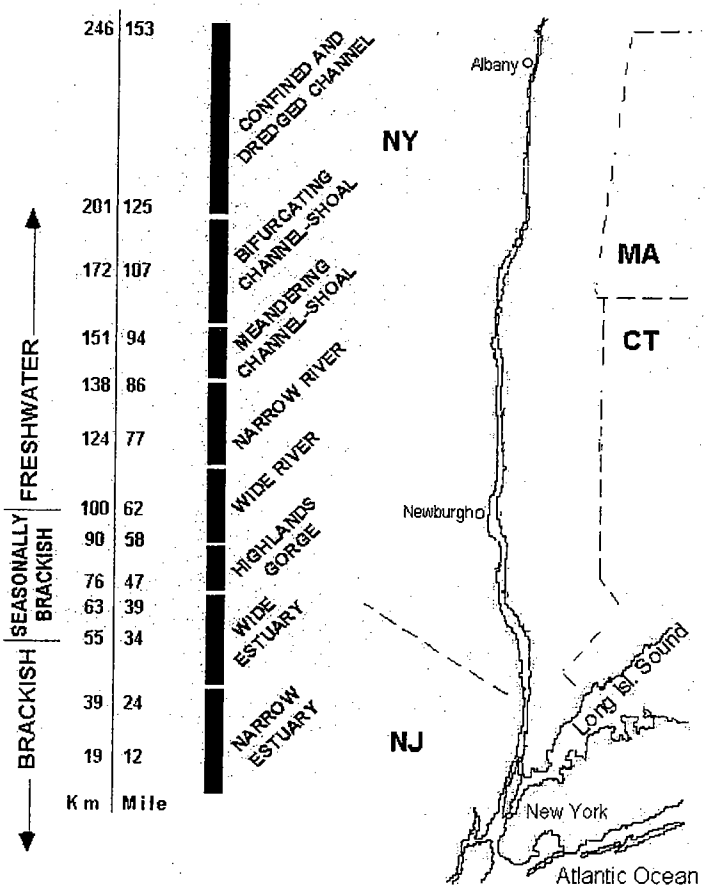
located in the lower St. Lawrence River prior to their discovery in Lake Ontario in 1873. From Lake Ontario, alewives are thought to have invaded the lower Great Lakes via the Lake Erie Barge Canal to Lake Erie, where they were first observed in 1933 (Ihssen et al. 1992). Alewives were then subsequently reported in Lake Huron, Lake Michigan, and Lake Superior in 1933, 1949, and 1954, respectively (Ihssen et al. 1992).

This study was aimed at characterizing microsatellite variation in alewife populations in the Hudson River. Microsatellites are currently the most common type of marker used in genetic analyses. Microsatellites are short, tandem, nucleotide repeats that are considered to be neutral and free from selective pressures. Microsatellites are utilized because they are very abundant and spread over the entire genome of all living organisms, so markers can be readily developed for any genetic objective. Microsatellites often exhibit extremely high levels of allelic variation. The main objective of the research was to determine the extent of genetic variation among different alewife populations in the Hudson River, and to compare that variation to other anadromous and freshwater populations.

METHODOLOGY

Alewives spawn in a number of tributaries in the Hudson River. Alewife tissue samples were collected in the summer of 2003 from young-of-the-year (YOY) alewives at two locations in the Hudson River. Individuals at the Hudson River Estuary Program were kind enough to collect samples for our research. Samples were collected from Newburgh at the beginning of the YOY migration to sea (early July) and from Albany during the middle to end of the migration to sea (late August) (Figure 1).

Figure 1: Map of the Hudson River.



DNA was extracted from muscle tissue preserved in 95% EtOH using the tissue protocol specified by Qiagen's QIAamp® DNA Mini Kit (Qiagen Inc., Valencia, CA). Previously collected alewife populations from locations other than the Hudson River are included here for comparison. The anadromous locations were the Union River, ME; the St. Croix River, ME; Wight Pond, ME; the Matapedia River, Canada; Hollyroud, Newfoundland; and the St. Lawrence River; the freshwater population was from Seneca Lake, NY. Microsatellites were used to characterize all populations of alewives. Microsatellites used in this research were developed by Brown et al. (2000) initially for

use with American shad (*Alosa sapidissima*) and were adapted for use in this study. The six loci examined were: Asa-2, Asa-4, Asa-8, Asa-9, Asa-12 and Asa-16 (Table 1).

Table 1. Microsatellites primers used to characterize alewife (*Alosa pseudoharengus*) populations.

Locus	GenBank Accession	Repeat Motif	Size Range (bp)	Number of Alleles	Annealing Temperature °C
Asa-2	AF039657	$(TTC)_{13}$	79-115	11	50
Asa-4	AF039658	$(ACC)_2(AAC)_{12}(AGC)_6$	106-142	13	48
Asa-8	AF039660	$(TTTG)_8$	108-148	9	50
Asa-9	AF039661	$(TTTC)_7$	157-309	38	50
Asa-12	AF039663		334-390	14	52
Asa-16	AF039662	$(GTT)_3(CCT)(GTT)_{12}$	101-128	10	52

The Polymerase Chain Reaction (PCR) enables the rapid amplification of DNA between two primers, or multiple sets of primers when multiplexing. This amplified DNA can then be assessed for sequence or size variability (Ward 2000). DNA amplifications were carried out in 25µl reactions volumes containing 1.5µl 25mM MgCl (Promega), 2.5µl 10x Thermophilic DNA Polymerase Buffer (Promega), 2µl 10mM dNTPs, 0.5µl 10mM forward primer, 0.5µl 10mM reverse primer (labeled at 5' with HEX, TET or 6-FAM fluorescent dye), 0.15µl *Taq* DNA Polymerase (Promega), and 2µl DNA template. PCR reactions were carried out in a PTC-100 programmable thermal cycler (MJ Research) as follows: an initial denaturation of 5 minutes at 95°C, followed by 35 cycles of denaturing at 95°C for 1 minute, annealing at 48-52°C for 1 minute and extending at 72°C for 2 minutes, with a final extension of 10 minute at 72°C. Asa-2, Asa-8 and Asa-9

were multiplexed in the same 25µl PCR reactions, while the remaining three loci (Asa-4, Asa-12, and Asa-16) were amplified individually. Fluorescent PCR products were visualized on an ABI377 automated DNA sequencer (Perkin Elmer) using an internal size standard (TAMRA) on all lanes. Fluorescent peak data were analyzed using GeneScan (Version 2.1) and Genotyper (Version 2.1) software programs (Perkin Elmer).

Samples from all populations were tested for conformation to Hardy-Weinberg equilibrium (HWE) expectations by the Markov chain method (Guo and Thompson 1992), resampling 10,000 iterations per batch for 200 batched in GENEPOP 3.1c (Raymond and Rousset 1995). Comparisons were made among and between populations for divergence of allelic distribution. An unbiased estimate of the value of a log-likelihood based exact test by the Markov chain method was calculated. Estimates of F_{st} were generated following Weir and Cocherham (1984). The extent of divergence among populations was quantified by the chord distance (D_{CE}) of Cavalli-Sforza and Edwards (1967) using GENETIX (Belkhir et al., 2002). D_{CE} measures genetic distance assuming no mutation with gene frequency changes created by genetic drift alone. This model does not assume that population sizes have remained constant and equal in all populations. Compared to other distance metrics, the use of D_{CE} leads to a higher probability of obtaining the correct tree topology under both the infinite alleles model (IAM) and the stepwise mutation model (SMM) (Takezaki and Nei 1996, Angers and Bernatchez 1998). Using neighbor-joining algorithms (Saitou and Nei 1987), pair-wise distances were employed to construct a population phenogram using MEGA version 2.1 (Sudhir et al. 2001). Confidence limits for the chord distances and F_{st} were estimated by the

percentage of 1000 bootstraps performed, resampling allelic frequencies using PHYLIP v. 3.5c (Felsenstein 1993).

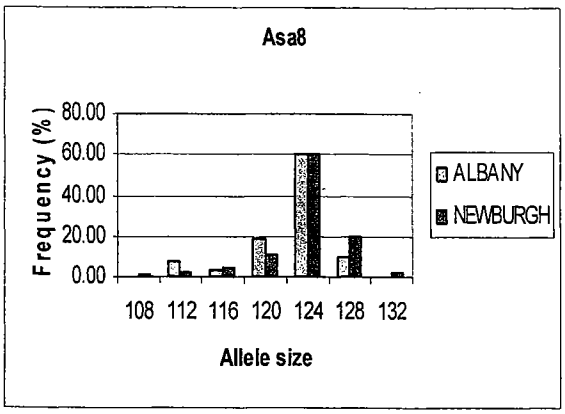
RESULTS

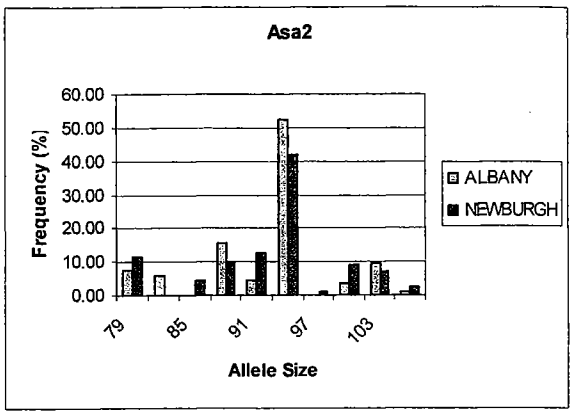
Samples were collected in July and August of 2003 by the Hudson River Estuary Program. A total of 108 alewives were examined from the Hudson River, with 58 samples collected from Albany, NY and 50 samples collected from Newburgh, NY (Table 2).

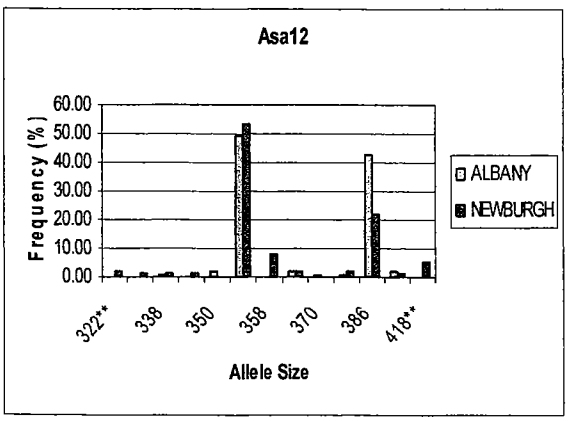
The exact tests of Hardy-Weinberg equilibrium (HWE) showed deficits in heterozygotes for both Hudson River populations. Two of the six microsatellite loci, Asa-4 and Asa-16, were out of HWE for both populations, while Asa-12 showed an additional deficit for the Newburgh population. None of the departures were significant.

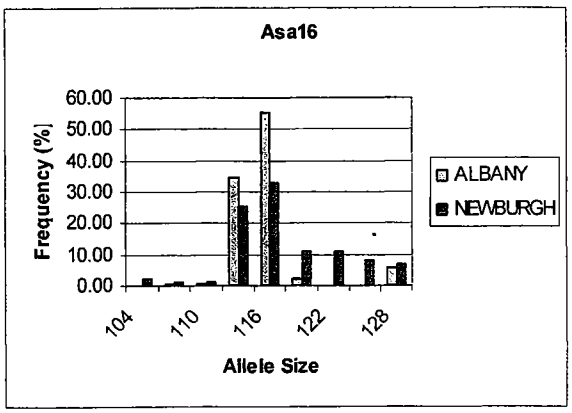
Each sample was genotyped at the six microsatellite loci and population allelic frequencies were determined (Figure 2). Exact tests of allelic differentiation were calculated for Albany and Newburgh at each locus and were significant for all loci. For Asa8 and Asa2, the allelic differentiation was significant with p-values of 0.041± 0.004 and 0.004±0.001 respectively. Additional significance was seen at Asa9, Asa16, Asa4, and Asa12 with p-values <0.001.

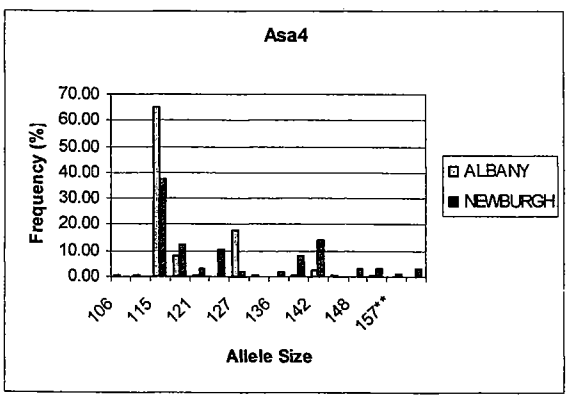
Figure 2- Allelic frequency distribution of the six microsatellite alleles at Albany and Newburgh, NY. **Represents alleles unique to the Hudson River

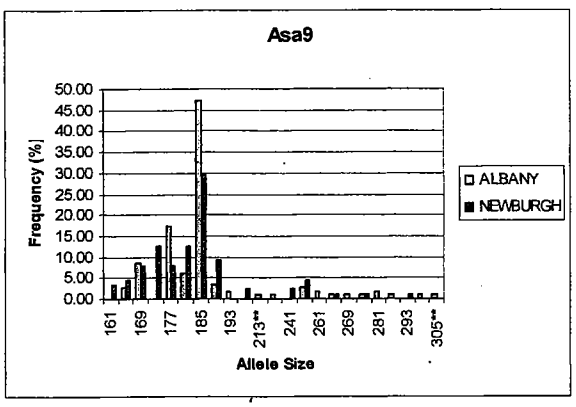












For anadromous populations, the expected (unbiased) heterozygosity values ranged from 0.525 to 0.750 (mean = 0.637) (Table 2). The Newburgh population had the highest value for expected heterozygosity at 0.750, while Albany population was below

the anadromous average at 0.618. Both populations have a large average number of alleles, with allelic richness values of 8.33 and 10.37 for Albany and Newburgh, respectively.

Table 2- Sample Size, Number of Loci, Unbiased Heterozygosity, Observed Heterozygosity and Average Number of Alleles for two Hudson River Populations and an Average of the Anadromous Populations

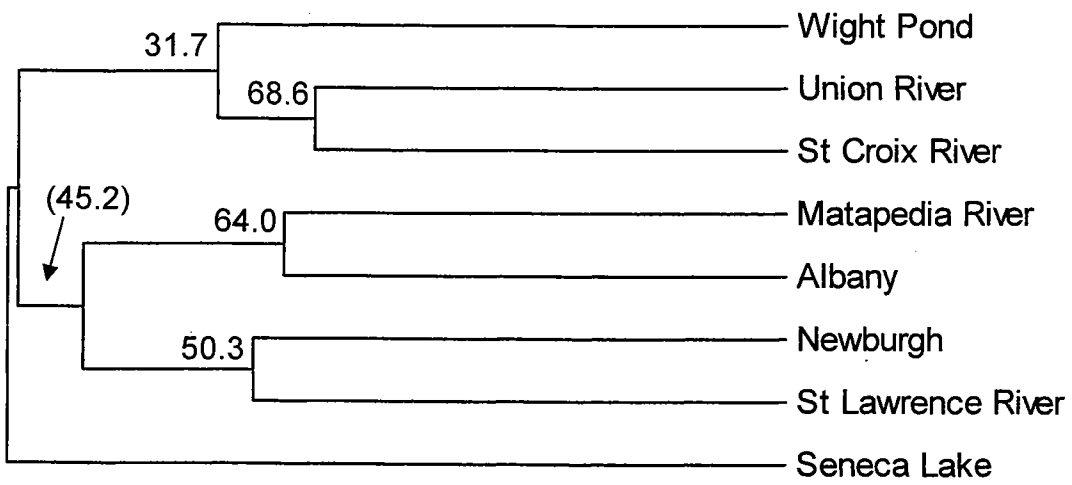
Population	Sample size	Loci typed	No Alleles	Unbiased Hz	Unbiased Hz SD	Obs Hz	Obs Hz SD
Albany	47.0	6	8.833	0.628	0.031	0.602	0.030
Newburgh	50.0	6	10.333	0.750	0.042	0.742	0.026
Average Anadromous	44.1	6	7.479	0.646	0.050	0.643	0.031

F_{st} has a range of 0 to 1, with 0 representing no genetic differentiation, and 1 representing complete genetic differentiation. For anadromous alewives, the values of F_{st} range from 0.023 – 0.206, with a mean value of 0.092. The two Hudson River populations, Albany and Newburgh, have an F_{st} value of 0.039 (Table 3). The chord distance between the Hudson populations is 0.033, while the range for all anadromous populations was from 0.019 – 0.091 (mean of 0.020). A Neighbor-joining tree was constructed from 1000 bootstrap replicates using chord distances for anadromous and freshwater populations (Figure 3). Based on microsatellite frequencies and chord distances, the Seneca Lake population branches outside of the anadromous populations, while the support between anadromous populations is low. This clustering appears artefactual with no bootstrap values above 68%, suggesting either that there is homogeneity among anadromous populations or that there is insufficient information in the microsatellite loci to establish relationships among populations.

Table 3: F_{st} comparisons between anadromous alewife populations (U= Union River, ME, SC= St. Croix River, ME, MR= Matapedia River, Canada, W= Wight Pond, ME, ALB= Albany, NY, NB= Newburgh, NY, LAW= St. Lawrence River, SL=Seneca Lake, NY).

	U	SC	MR	W	ALB	NB	LAW	SL
U	X							
SC	0.051	X						
MR	0.176	0.215	X					
W	0.062	0.072	0.120	X				
ALB	0.138	0.119	0.052	0.062	X			
NB	0.106	0.098	0.096	0.050	**0.039**	X		
LAW	0.147	0.130	0.115	0.088	0.072	0.040	X	
SL	0.192	0.147	0.231	0.109	0.132	0.105	0.068	X

Figure 3: Neighbor-joining tree using Cavalli-Sforza and Edwards chord distance to compare anadromous alewife populations and Seneca Lake. Numbers represent bootstrap support from 1000 replicates.



DISCUSSION

The aim of this study was to examine genetic variation of anadromous and freshwater alewives, and to determine if population structure exists in the Hudson River. Knowledge of evolutionary history and genetic relationships of fishes is essential to

monitoring stock structure and overall biodiversity of a species. The assessment of differentiation at neutral loci can represent a powerful tool to test if different forms are representative of alternative life histories in a single gene pool, or if they are due to reproductively isolated populations (Pigeon et al. 1997). Microsatellites are used extensively to assess the effective population size of stocks, stock identification, levels of inbreeding, population structure, gene flow, and parentage (Neff et al. 2000).

Anadromous alewife populations have, on average, higher levels of polymorphism than their freshwater counterparts (e.g., the Hudson River vs. Seneca Lake). For all loci, the Hudson River populations shared a number of alleles that were common to all anadromous populations. However, the Hudson populations did show a number of unique alleles. The presence of a large number of alleles, compared to other anadromous and freshwater populations, suggests that population mixing may be occurring. If true, there may be genetic population structuring in the Hudson River.

For the two Hudson populations, the excess number of alleles has resulted in deviations from Hardy-Weinberg equilibrium (HWE), primarily in the form of heterozygote deficiencies, suggesting a possible Wahlund Effect (Hartl and Clark 1997). In the Wahlund Effect, subpopulations with different allelic frequencies mix, and heterozygote deficiencies occur. Following random mating, HWE is restored. However, in this case, there is no reason to believe that random mating among YOY would naturally occur. The presence of significantly different allelic distributions in the two populations also suggests that population structure is present in the Hudson River.

Both F_{st} (0.039) and Cavalli-Sforza and Edwards chord (0.033) distances between the two Hudson River populations is consistent with differentiation within the Hudson,

although the levels are small when compared to those of other anadromous populations (means of 0.092 and 0.020, respectively). The topology the neighbor-joining tree suggests that differentiation exists between anadromous alewife populations, but the data are insufficient to demonstrate differentiation in the Hudson River: Albany and Newburgh do not cluster together and bootstrap values are low. Seneca Lake did not cluster within anadromous populations, suggesting genetic differentiation between the two habitat types. Overall, however, there is significant genetic differentiation between the two Hudson River alewife populations as demonstrated via the exact tests of allelic differentiation.

CONCLUSIONS

The level of genetic differentiation between population samples in the Hudson River is relatively low when compared to that of other anadromous populations. To examine the question further, I suggest looking at adult alewives at the beginning of the spawning runs in different tributaries of the Hudson River. This would provide better representation of the structure of Hudson River alewife populations.

ACKNOWLEDGEMENTS

I would like to thank the Hudson River Foundation and the Association for Graduate Students at the University of Maine for their financial support of my Masters research. I would like to thank Bonnie Brown for the use of her microsatellite primers and to Kathy Hattala, Amanda Cosman, and Jennifer Temple of the Hudson River Estuary Program for providing alewife samples.

REFERENCES

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

Bean, T. 1884. On the occurrence of the branch alewife in certain lakes of New York. *In* Goode, G. B. The fisheries and fishing industries of the United States. Section 1, 588-593.

Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2002. GENETIX 4.04, logiciel sous Windows TM pour la genetique des populations. Laboratorie Genome, populations, Interactions, CNRS UMR 5000, Universite de Montpellier II, Montpellier (France).

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

Cavalli-Sforza, L. and A. Edwards. 1967. Phylogenetic analysis: models and estimation procedures. Evolution 32:550-570.

Daniels, R.A. 2001. Untested assumptions: the role of canals in the dispersal of sea lamprey, alewife and other fishes of the eastern United States. Environmental Biology of Fishes 60: 309-329.

Felsenstein, J. 1993. *PHYLIP (Phylogeny inference package), Version 3.5c*. Department of Genetics, SK-50, University of Washington, Seattle, WA.

Gnaedinger, R. and R. Krzeczkowski. 1966. Heat inactivation of thiaminase in whole fish. Commercial Fisheries Review 28 (8): 11-14.

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

Hartl, D. and A. Clark. 1997. Principles of Population Genetics. Sinauer Associates, Inc. Sunderland, MA. 3rd Edition.

Ihssen, P.E., G.W. Martin and D.W. Rodgers. 1992. Allozyme variation of Great Lakes alewife, *Alosa pseudoharengus*: genetic differentiation and affinities of a recent invader. Canadian Journal of Fisheries and Aquatic Sciences 49:1770-1777.

Ketola, H.G., P. Bowser, G. Wooster, L. Wedge and S. Hurst. 1999. Effects of thiamine on reproduction of Atlantic salmon and a new hypothesis for their extirpation in Lake Ontario. Transactions of the American Fisheries Society 129: 607-612.

Neff, B.D., P. Fu and M. Gross. 2000. Microsatellite multiplexing in fish. Transactions of the American Fisheries Society 129:584-593.

Neilands, J. 1947. Thiaminase in aquatic animals of Nova Scotia. Journal of the Fisheries Research Board of Canada 7: 94-99.

Neves, R. 1981. Offshore distribution of the alewife, *Alosa pseudoharengus*, and the blueback herring, *Alosa aestivalis*, along the Atlantic coast. Fishery Bulletin 79:473-485.

Pigeon, D., A. Chouinard and L. Bernatchez. 1997. Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). Evolution 51:196-205.

Raymond, M. and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenism. Journal of Heredity 86:248-249.

Saitou, N. and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425.

Smith, H. 1890. Report on an investigation of the fisheries in Lake Ontario. House of Representatives. Miscellaneous Document 341, and Bulletin of the U.S. Fish Commission 10: 177-215.

Smith, S. 1995. Early changes in the fish community of Lake Ontario. Great Lakes Fishery commission. Technical Report 60. Ann Arbor, Michigan.

Sudhir K., K. Tamura, I. Jakobsen, and M. Nei. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software, Arizona State University, Tempe, Arizona, USA.

Takezaki, N. and M. Nei. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144:389-399.

Ward, R.D. 2000. Genetics in fisheries management. Hydrobiologia 420:191-201.

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

Wright, R. 1892. Preliminary report on the fish and fisheries of Ontario. *In* Ontario Game Fish Comm. Report P. 419-475.