INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road Ann Arbor MI 48106-1346 USA
313 761-4700 800 521-0600

Order Number 9304710

Mitochondrial DNA analysis of American shad (*Alosa sapidissima*) populations

Nolan, Kathleen Ann, Ph.D.

City University of New York, 1992

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

MITOCHONDRIAL DNA ANALYSIS OF AMERICAN SHAD (ALOSA SAPIDISSIMA) POPULATIONS

by

Kathleen A. Nolan

A dissertation submitted to the Graduate Faculty in Biology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1992

This manuscript has been read and accepted for the Graduate Faculty in Biology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 27, 1992
date

Chairman of Examining Committee
Dr. Sharon Cosloy, City College

JULY 29, 1992
date

Executive Officer
Dr. Peter C. Chabora

Dr. Jane Gallagher, City College

Dr. Joseph Rachlin, Lehman College

Dr. Isaac Wirgin, New York University

Dr. Robert Schmidt, Simons Rock College

Dr. C. Lavett Smith, American Museum of Natural History

Supervisory Committee

The City University of New York

Abstract

MITOCHONDRIAL DNA ANALYSIS OF AMERICAN SHAD (ALOSA SAPIDISSIMA) POPULATIONS

by

Kathleen A. Nolan

Advisors: Drs. Joseph Grossfield and Sharon Cosloy

Restriction endonuclease analysis of highly purified mitochondrial DNA (mtDNA) was used to differentiate among spawning stocks of American shad (Alosa sapidissima), in nine major Western Atlantic rivers. Four and five base-cutting restriction enzymes, as well as one six base-cutting enzyme, were used to prepare both individual enzyme profiles and composite genotypes. Most of the differences observed in mtDNA types were attributed to small length polymorphisms visible on polyacrylamide gels. Significant differences among frequency distributions of mtDNA genotypes generated by single restriction enzyme digests were detected among all rivers. Two stock-specific markers were detected in moderate frequencies in two rivers: a base substitution on EcoR I agarose gels (27% in the St. Johns River in Florida) and a length variant on Dde I polyacrylamide gels (62% in the Miramichi River in New Brunswick, Canada). One genotype revealed on Dde I agarose gels decreased in frequency with latitude.

Ninety-eight composite mtDNA genotypes were detected out of 161 fish tested. Forty-six percent of the genotypes were found in only single fish. Divergence values among composite genotypes were low (0.00057-0.0042)

and similar to those of other anadromous fish. The divergence between American shad and hickory shad (0.055-0.0152) was an order of magnitude less than expected on the basis of a molecular clock calibrated against the fossil record.

Phenograms and cladograms were calculated which related shared composite genotypes as unrooted as well as trees rooted using hickory shad as an outgroup. The resultant analyses do not support an hypothesis of simple south to north colonization of rivers as a result of glacial retreat. The complexities of the trees may be the result of infidelity in homing as well as stocking of rivers in the recent past. The complexities of the trees and the diversity in the rivers suggest that the mtDNA analyses conducted in this study cannot be used reliably to identify river-specific stocks of shad in an open ocean fishery for the Western Atlantic using currently available methodology.

Acknowledgements

I would like to remember Joseph Grossfield as being a wonderful mentor and a good and loyal friend. He was always very patient and encouraging and was an excellent teacher. He showed me how to do the techniques required for the project with great precision and care.

I thank Isaac Wirgin for his constant attention to the project, for helping me become proficient at the techniques that I learned and continual follow-up and guidance.

I am grateful to Jane Gallagher for help with some of the computer programs and for getting me through the revisions.

Paul Margolin was a source of inspiration and help throughout most of the course of preparing this dissertation.

My committee members gave me lots of advise and help in preparing this project and I thank them for. it.

I would like to thank the people at City College for being so helpful.

I thank all the people who helped me to collect and isolate the mtDNA from the shad, especially Rui Proenca.

Nick Hoffman was invaluable in helping me learn the Macintosh.

Katherine Keenan and Barry Potvin are remembered for encouraging me to study for the doctorate.

I especially thank my husband, Richard Conley for accompanying me on long collecting trips, and putting up with the late nights at the lab and at the computer.

This research was supported by grants 666281 and 667145 to Joseph Grossfield from the PSC-CUNY Research Award Program of the City University of New York.

TABLE OF CONTENTS

List of Tables — Page No.

List of Figures — Page no.

INTRODUCTION

Migration and homing

The American shad (Alosa sapidissima) is an anadromous fish of considerable economic importance throughout its western Atlantic range, which extends from the St. Johns River in northern Florida to the St. Lawrence River in Canada (Walburg and Nichols 1967; Scott and Crossman 1973). Historically, spawning runs have been documented in at least 23 rivers within this range (Walburg and Nichols 1967). Shad spawn in the fresh water portions of the rivers. Immature fish spend the winter along the mid-Atlantic coast, migrate northward to Gulf of Maine with the spawned adults in the spring, and return to the mid-Atlantic coast in the late fall (Talbot and Sykes 1958).. Shad spend their post-juvenile lives (4-6 years) in the ocean before they return to their natal rivers in the spring to spawn (Talbot and Sykes 1958; Neves and Depres 1979). In the warmer months, adult fish are found in the Bay of Fundy, the St. Lawrence estuary and off the Labrador coast. Mark-recapture studies of fish tagged during the summer in the Bay of Fundy, the St. Lawrence estuary, and off the Newfoundland and Labrador coasts have framed the overall pattern of shad migration and have provided recaptures from all the major shad spawning rivers along the Atlantic coast of North America (Dadswell et al. 1987).

Mixing of juvenile and adult shad representatives of diverse spawning systems apparently occurs in the coastal waters extending from Georgia to the Bay of Fundy, the St. Lawrence estuary, and off the Newfoundland and Labrador coasts (Dadswell et al. 1987) It is unclear whether the representatives of the various stocks migrate synchronously in time and space. Until stock identification of shad taken from a mixed ocean sample is feasible, the regional composition and extent of offshore mixing cannot be documented (Neves and

(Neves and Depres 1979).

Mark-recapture studies have shown that American shad are relatively faithful homers (about 97%) (Hollis 1948; Mansueti and Kolb 1953; Williams and Dabom 1984; Dodson and Leggett 1973; Melvin et al. 1986). Meristic data have suggested that shad not only home to their natal rivers but also to tributaries within the St. John River, New Brunswick (Carscadden and Leggett 1975). However, U. S. Fish Commission reports (1871-1886) show that extensive shad and stocking programs were conducted along the U. S. eastern coast. In come cases, eggs, juveniles, and adult fish were transplanted to non-natal rivers. This could have been a source of "artificial infidelity" that could complicate stock identification.

Variation in population abundance

Commercial landings of shad have declined over the past 100 years. Annual catches have dropped from a record high of 50 million pounds in 1897 to current levels approaching 2-3 million pounds (Mansueti and Kolb 1953; Richkus and DiNardo 1984). The past two decades especially have witnessed a precipitous decline in commercial landings of this species, possibly indicative of a reduced population size.

While the overall catch for the Atlantic seaboard has decreased, each river system has its own pattern of shad abundance (for a review of American shad catch statistics see Cooper (1984) and references therein). Catches in the three major northern breeding populations in the United States (in the Connecticut, Hudson and Delaware rivers) are stable or even increasing in number (Brandt 1987; Gibson et al. 1988; Richkus and DiNardo 1984). This could be due to ameliorating factors such as the construction of a fish lift at the Holyoke dam on the Connecticut River in 1955 and the reduction of a pollution block at the

mouth of the lower Delaware River. In contrast, rivers south of Delaware Bay have witnessed a steep reduction in population size (Gibson et al. 1988; Richkus and DiNardo 1984; Williams and Bruger 1972). The complete closure in 1980 of the once lucrative Maryland shad fishery has been one reaction (Richkus and DiNardo 1984).

Contrasting life history strategies (Leggett and Carscadden 1978) and recruitment mechanisms (Summers and Rose 1987; Savoy and Crecco 1988) as well as pollution, hydrographic and meteorological conditions, fishing mortality and man-made obstructions may contribute to the disparate health of the various shad breeding populations.

Contrasting life history strategies

There are no repeat spawners south of the Neuse River, North Carolina or 35 N (Leggett and Carscadden, 1978). Above this latitude, the proportion of repeat spawners increases directly with latitude, from approximately 25% in the Potomac River to as high as 80% in the St John River in New Brunswick, Canada. There is a concordant increase in size with increasing latitude and a linear decrease in fecundity of shad with latitude. Gibson et al. (1988) also point out that life-history traits such as higher fecundities, shorter life spans and earlier maturation times are found in species that live in more stable environments, as opposed to unstable northern environments.

Recruitment variation

Shad year-class strength is significantly correlated with density independent factors such as mean river-discharge, water temperatures and total monthly precipitation in the Connecticut River (Crecco and Savoy 1984; Savoy and Crecco 1988). A percentage of eggs and larvae mortality is

attributable to density-dependent factors, such as intraspecific competition for food and space and/or predation. Density-dependent factors are also affected by density-independent factors, i. e. the amount of space available can be influenced by rate of river flow (Crecco and Savoy 1987; Savoy and Crecco 1988). There is no correlation between numbers of spawning adults and future recruitment into the adult population (parent-progeny relationship) in the Connecticut River based on population data from 1966-1982 (Crecco and Savoy 1985). The same authors (Savoy and Crecco 1988) found a positive correlation between the abundance of early larvae and all subsequent life history stages of shad in population data from 1979 through 1987 and adult recruitment from the relative year classes. Apparently more favorable environmental conditions existed for shad in those years. Gibson et al. (1988) also noted that there was a lower variability of flow rates near the center of the shad's range which is a density-independent parameter that affects abundance in shad populations.

Effects of pollution on shad populations

Summers and Rose (1987) found that shad stock abundance in the Potomac, Hudson and Delaware rivers was related to sewage (Potomac) loading and dissolved oxygen levels (Hudson and Delaware) as well as hydrographic conditions such as river flow and water temperature. Stock abundance declines with increasing sewage loading and biological oxygen demand and decreasing dissolved oxygen levels.

Effects of fishing on shad populations

Shad stocks in the southern rivers are thought to be depleted through over-fishing Gibson et al. (1988). Harvest of the mixed ocean fishery occurs at

several different locales and seasons. During the late fall and winter this resource is vulnerable to an international offshore midwater trawl fishery of unknown magnitude (Richkus and DiNardo 1984). In addition, shad of presumed mixed origin are also harvested during their spring northerly spawning migration by nearshore fisheries, particularly in southern states (Richkus and DiNardo 1984). For example, the ratio of ocean to inshore landings was greater than ever before for South Carolina in 1982. The increase in South Carolina offshore landings may have actually masked a decline of its resident stock, since the geographic origin of the ocean fish is not known. (Richkus and DiNardo 1984). Assignment of river of origin to these ocean fish would clarify the problem of their ancestry.

Hydroelectric tidal power plants

Future mortality for this mixed stock may also occur as a result of potential operation of Canadian hydroelectric tidal power plants at known critical summer foraging and spawning grounds in the Bay of Fundy (Williams and Daborn 1984; Dadswell et al. 1986).

Management options

A. Pollution reduction

River-specific actions to reduce pollution would be a viable shad management option. A few examples of pollution reduction on individual rivers are as follows: A law was passed in the state of Maryland in 1985 banning phosphates in detergents in order to limit the amount of phosphates flowing into rivers such as the Potomac and Patuxent rivers (Barth 1989). Sewage treatment plants have been built or upgraded on these rivers. As a result of these measures, the dissolved oxygen levels in these rivers has increased. The

North River sewage treatment plant on the Hudson River began operations several years ago in an attempt to lower biological oxygen demand and increase dissolved oxygen levels.

B. Fish passageways

Other river-specific management strategies include building fish passageways over dams to allow shad to migrate to spawning runs. These have been successful on the Connecticut River (Richkus and DiNardo 1984) and are being built on several tributaries of the Chesapeake (Chesapeake Bay Foundation 1989).

C. Fishing options

Gibson et al. (1988) found that there was a parabolic relationship between the riverine maximum sustainable fishing rates and latitude. Their findings suggest that northern (Pawtucket and Connecticut rivers) and southern (St. Johns and Altamaha rivers) American shad stocks are less able to recover from high fishing mortality rates (greater than 50%) than are stocks (excepting depleted ones such as the Susquehanna River) in the middle of the shad's range. They suggest that maximum harvest rates be kept at 50% or below, particularly for the northern and southern rivers afore-mentioned.

Importance of discrimination of shad populations

Fishing mortality in the open ocean and at inshore coastal sites may potentially negate the beneficial effects of management procedures on the individual spawning rivers. A means of identifying ancestry of fish in this mixed oceanic stock is needed as a tool which could be used to protect threatened shad breeding populations. Stock discrimination was identified as an area of highest priority by the workshop on critical data needs for shad research on the Atlantic coast of North America (Cooper 1984)

Discrimination of shad populations would allow for the identification of both the river of origin of fish in mixed stocks as well as estimates of contribution of individual river systems to the overall Atlantic fishery. It would also permit assessment of migration routes and the detection of movements of unique contingents in the overall annual migrations. At a more detailed level, population-specific tags would be useful in assessing differential utilization of feeding resources as well as relative survival with respect to environmental variables. If an identification method is applicable to fish of various ages, this would provide important information with respect to the degree of population cohesion at various points in their life histories.

Traditional approaches to stock identification of shad

Several traditional approaches have attempted to discriminate shad stocks. Hudson River shad have sometimes been divided into two races called "bluebacks" and "yellowbacks" based on mostly qualitative characteristics such as coloration (Greeley 1936).

Meristic studies reveal significant differences among shad representative of the Miramichi, Connecticut, Hudson and St John (New Brunswick) rivers and among tributaries and the main body of the St John River. Univariate and multivariate analyses of five characters reveal significant differences in mean numbers of meristic characters among rivers and among St John River tributary comparisons (Carscadden and Leggett 1975).

Melvin (1984) studied 10 meristic and 16 morphometric characters of shad collected from 13 rivers throughout the shad's range. He developed linear discriminate functions that were used to assign origin of shad caught in the Cumberland Basin. during the summer months to four specific regions. These were Cape Hatteras and south, Cape Cod to Cape Hatteras, the Bay of Fundy

and the Gulf of St. Lawrence.

Another method used was analysis of zonation patterns in otoliths (Williams 1985). This technique was used to correctly assign (93% of the time) the stock origin of shad from the mixed summer fishery in the Cumberland Basin to three broad geographic groups, encompassing Canadian Atlantic, Mid-U.S. Atlantic, and South-U.S. Atlantic rivers.

However, the identification of origin of fish by meristic and morphometric analysis can be difficult due to overlap in individual counts, often limiting the applicability of such analyses to the question of relative contribution to mixed fisheries.

Although meristic and morphometric characters have a genetic component to their expression, they are also subject to temporal instability due to environmental influences (Blaxter 1984), thereby potentially requiring the annual recalibration of discriminators. For example, significant temporal variability in meristic counts among year classes of several species of alosids was reported in tributaries of the Chesapeake Bay (DesFosse and Loesch 1985), although Carscadden and Leggett (1975) found little difference in meristic counts among year classes of shad in the St. John River, New Brunswick. The latter authors think the numbers of meristics are genetically controlled, but suggest that environmental changes may influence numbers of meristic characters over time. A correlation between environmental variables and meristics was not attempted by this study.

Genetic approaches to shad population discrimination

A. Protein electrophoresis

Genetic approaches might eliminate the need for annual recalibration of discriminators. However, electrophoretic analysis of proteins does not reveal

significant genetic heterogeneity among shad collected at several sites in the Chesapeake Bay and the Nansemond River (Sismour and Birdsong 1986). Of seven enzymes studied, two are polymorphic. A high degree of homozygosity is associated with the Nansemond. This is consistent with findings with other anadromous fish species, such as striped bass from the Chesapeake (Sidell et al. 1980).

B. Mitochondrial DNA analysis

In the present study, restriction fragment length polymorphisms (RFLP) analysis of mitochondrial DNA (mtDNA) was used as a genetic tag in an attempt to discriminate populations of shad. MtDNA is a double-stranded, closed-circular, superhelical molecule of relatively small size in vertebrate animals, from 15,700 to 19,500 base pairs (bp), (Densmore et al. 1985) found in multiple cellular copies. Its base sequence, and in some forms, its genetic code is different from that found in nuclear DNA (Anderson et al.1981). The genetic map of a fish (carp--Cyprinus carpio) mtDNA molecule was initially described by Araya et al. (1984). The carp molecule shows conservation of organization with the human mtDNA genome, which codes for 2 rRNA and 22 tRNA genes and 13 genes that code for mitochondrial protein products, primarily utilized in electron transport and oxidative phosphorylation (Anderson et al.1981). The Cyprinus carpio molecule is 16,400 base pairs long (bp) which is similar in size to that of other vertebrates (Anderson et al. 1981; Densmore et al. 1985). MtDNA sizes in other fish species ranges from 16300-19800 base pairs (Avise et al. 1984; Berg and Ferris 1984; Graves et al. 1984; Bentzen et al. 1988; Graves and Dizon 1989.

The size and conformation of mtDNA allows for its fairly simple isolation, and mtDNA apparently evolves at a rate ten times that of the nuclear genes underlying structural protein and meristic markers (Brown et al. 1979). MtDNA

is usually transmitted maternally (Lansman et al.1983), and, in the absence of further mutation, all progeny of a single female share identical mtDNA genotypes. This permits an assignment of maternal ancestry of individual fish although no information is generated concerning paternal lineage. The relatively rapid mutation rate (but relatively conserved gene order) lack of recombination (Gyllensten et al. 1985) and repair mechanisms (Clayton et al. 1974) allow it to be treated as a haploid molecule, fixed changes in which can be followed from maternal generation to generation.

Mitochondrial DNA may be isolated and purified from individual fish of interest. Isolated mtDNA molecules may be cleaved at specific oligonucleotide recognition sites by commercially available enzymes termed Class II restriction endonucleases. Each restriction endonuclease recognizes a very specific four to eight base pair sequence in double-stranded DNA molecules. This recognition site varies among enzymes; more than 100 enzymes are now commercially available.

Cleavage of a circular molecule of mtDNA by one or several enzymes in concert reduces the mtDNA to a collection of discrete fragments. The number of fragments will equal the number of recognition sites for the enzyme used. The number of recognition sites cannot be predicted, but may be estimated statistically. For instance, assuming DNA sequences are randomly distributed and all four nucleotides are present equally, a six-base recognition site will be found with a frequency of $(1/4)^6 = 2.4 \times 10^{-4}$, or every 4096 bases (Nei and Li 1979). Thus, four fragments are predicted for metazoan mtDNAs, and a larger number (approximately 30) is predicted for 4- and 5-base cutting restriction endonucleases.

Unique mtDNA genotypes representative of different spawning systems might be expected because of the rapid rate of mtDNA mutations and

population fixation. However, marine and anadromous fish populations might have a lower percentage of unique genotypes than fresh-water and terrestrial vertebrate populations because the non-oceanic populations may experience more barriers to gene flow.

MtDNA studies of other vertebrates

Population-specific fixed mtDNA genotypes are frequently observed in terrestrial animals, freshwater fishes, and estuarine fishes, and in marine and diadromous fishes with varying degrees of success of population discrimination. Most of these studies use six-base (and sometimes a few five-base) cutting restriction enzymes. The trend in recent mtDNA studies is not only to describe differences in mtDNA genotypes that exist in populations but to ascertain if there is concordance of genotypes with geography. Avise (1989) coined the phrase "phylogeography" for the study of systematics as related to biogeography. Phylogenies can "reflect" what happened geologically to a geographic area, i. e. if there is a population that harbors distinct mtDNA genotypes not contained in an adjacent population, perhaps a barrier to gene flow exists now or did in the past.

Lamb et al. (1989) observed a distinct genetic divergence in mtDNA genotypes between eastern and western populations of desert tortoises (Xerobates agassizi) along the banks of the Colorado River. The Colorado itself could have served as a barrier to dispersal of these populations, or the history of marine incursions and retreats may have further limited any contact between the two populations.

MtDNA studies of other fishes

The freshwater sunfish (Lepomis punctatus, L. gulosus, L. microlophus), and bowfin (Amia calva) from rivers in the southeastern United States exhibit distinct east-west mtDNA genotypes (Bermingham and Avise 1986). These breaks appear to be congruent with geologic events during the Pliocene which caused fragmentation of populations.

Gonzalez-Villasenor and Powers (1990) used mtDNA analysis to try to distinguish between primary and secondary intergradations in Fundulus heteroclitus populations. Primary intergradations are genetic changes that occur along a cline as a result of a difference in habitat. Secondary intergradations refer to genetic changes that occur along a cline that was once discontinuous because of geologic barriers. They found a northern population that includes Fundulus from Maine and one collecting site in New Jersey, and a southern population that includes fish from a different collecting site (more Southern) in New Jersey and Georgia. A disjunction occurrs between the two New Jersey sites which supports the hypothesis of secondary intergradation.

Graves et al. (1984) and Graves and Dizon (1989) could not distinguish Atlantic and Pacific populations of skipjack tuna (Katsuwonus pelamis) and albacore tuna (Thunnuatus alalunga). This is attributed to either recent isolation of the Atlantic and Pacific populations, migrations between the two oceans, or a small sample size, in which case, it was possible that differential markers were not detected.

RFLP analysis of mtDNA has also been used in an effort to discriminate populations of marine fishes, but usually with more ambiguous results than with terrestrial vertebrates, fresh-water and estuarine fish. Avise et al. (1987) attempted to discriminate populations of mouthbrooding marine catfish (Arius felis and Bagre marinus) and demersal toadfish (Opsanus beta and Opsanus

tau). Populations of toadfish are more readily discriminated perhaps because their mobility is more limited than that of the catfish. Kornfield and Bogdanowicz (1987) sought to determine if RFLP mtDNA genotypes were concordant with specific oceanic spawning sites of the Atlantic herring (Clupea harengus). A large number of mtDNA composite genotypes (a "composite" genotype is a genotype that is composed of more than one single enzyme genotype) generated--26 out of 69 fish tested--is attributed to a large population size. Some mtDNA genotypes are common to all spawning areas and some are rare (found in only one sample). There are no spawning area-specific mtDNA genotypes that occur more than once. Avise et al. (1989) also invoked the large population size (hundreds of million) of Atlantic menhaden (Brevoortia tyrannus and Brevoortia patronus) from the Gulf and Atlantic coasts as contributing to the large number of composite mtDNA genotypes found: 32 different genotypes found in 33 fish sampled. Two populations of Atlantic cod (Gadus morhua) from what are thought to be discrete spawning sites, one from the North Sea and the other from the Grand Banks of Newfoundland do not reveal major differences in mtDNA genotypes (Smith et al. 1989). Atlantic tomcod (Microgadus tomcod) (Wirgin et al. 1990a) also do not display variations in mtDNA genotypes. The catadromous eel populations (Anguilla anguilla) do not reveal any discrete mtDNA variations but this species does display genotypes that differ significantly from the European Anguilla rostrata (Avise et al. 1986). The composition of Anguilla anguilla populations is believed to be of mixed origin in the ocean, with widespread dispersal of larvae which would explain the lack of partitioning of mtDNA genotypes.

MtDNA restriction enzyme assays of seven Pacific salmon (Oncorhynchus tshawytscha) populations from British Columbia and Alaska fail to reveal discrete stock polymorphisms. Of the six genotypes uncovered, four are unique

and three are shared in fish from more than one population.

Davidson et al. (1989) (review paper) reveals that workers using six-base cutting restriction endonucleases can not discriminate anadromous Atlantic salmon (Salmo salar) populations in Newfoundland (Birt et al. 1986) Finland (Palva 1986) or the Baltic or North Sea (Gyllensten and Wilson 1987). However, Palva (1986) could distinguish anadromous Finnish stocks with four- and five-base-cutting restriction enzymes.

Polymorphisms in striped bass (Morone saxatilis) populations are attributable to base substitutions and length polymorphisms attributed to unknown mechanisms (Chapman, 1987, 1989, 1990; Wirgin et al. 1989, 1990b). Some polymorphisms are branded as heteroplasmy, which is two or more forms of the mtDNA molecule in the same fish. Both sets of workers found relatively few polymorphisms among striped bass populations. There are some groupings of mtDNA that are concordant with geographic locations within the Chesapeake. Wirgin (1987) attributes the low diversity within the striped bass populations to recent separation of populations.

Bernatchez et al. (1989) found less mtDNA variability in the anadromous North American whitefish (Coregonus clupeaformis) populations from the James and Hudson Bays than in anadromous European whitefish (Coregonus lavaretus) from the Bothnian Gulf. Population bottlenecks are implicated as being the cause of the lack of mtDNA diversity in the North American whitefish; mtDNA analysis could not distinguish specific stocks in either group. Bernatchez and Dodson (1990) did find frequency distributions of mtDNA genotypes that were concordant with discrete anadromous stocks of the lake cisco (Coregonus artedii) in the Hudson Bay, but not in the James Bay.

Bentzen et al (1988, 1989) distinguished three clusters of American shad (Alosa sapidissima) that agreed with overwintering geographic groupings of

American shad (Dadswell et al. 1987).

MtDNA isolated from American shad from nine rivers was digested with several restriction endonucleases in an attempt to answer the following questions:

1. Can restriction endonuclease analysis of mtDNA be unequivocally used to discriminate shad populations? This could be revealed by significant differences in genotypic frequencies between populations or geographic areas, or by unique genotypes that differentiate all fish.
2. Are the populations sufficiently differentiated so that if, at a future date, a shad of unknown origin was captured from a mixed ocean sample it might be assigned to its river of origin?
3. How heterogeneous and genetically divergent were the populations within and among each river?
4. Is more resolution or discriminatory power gained with the use of four- and five- cutting restriction endonucleases than with six-base cutting enzymes used in a previous study on American shad (Bentzen et al. 1988; 1989)?
5. How divergent are the American shad from the hickory shad (Alosa mediocris)?
6. How do American shad divergence values compare with those of other marine and diadromous species?
7. How are the American shad populations related phylogenetically to each other?
8. What is the best method for determining the above relationships?
8. Can fidelity of homing, dispersal patterns and gene flow be assessed from distinct geographic patterns in these genotypes?

9. How is this work similar to mtDNA studies done with other clupeids?

10. Can mtDNA analysis be applied in understanding problems in shad biology?

MATERIALS AND METHODS

Sample collections

Adult American shad were collected from 1984-7 during their spawning runs either personally or through contact with local fishermen or fisheries personnel from nine western Atlantic coast rivers: the St. Johns River, Florida (STJ), the Altamaha River, Georgia (ALT), the diversion canal between the Santee and the Cooper rivers, South Carolina (SAN), the Chowan River and the Albermarle Sound, North Carolina (CHO), the Rappahannock River, Virginia (RAP), the Delaware River, New York, New Jersey and Pennsylvania (DEL), the Hudson River, New York and New Jersey (HUD), the Connecticut River, Connecticut and Massachusetts (CON), and the Miramichi River, New Brunswick, Canada (MIR) (rivers and catch statistics and detailed descriptions of each sample are depicted in Table 1. (The above abbreviations for the nine rivers are used in all Tables and Figures.) (See Figure 1 for location of rivers.) A total of 316 American shad and one hickory shad were collected and subsequent procedures were carried out on all fish.

Ovaries or testes were excised from freshly killed fish. Some fresh shad ovaries or testes were shipped on ice via overnight Federal Express. Mitochondrial extraction procedures were initiated within two days of capture.

Mitochondrial Isolation

The protocol for mitochondrial isolation has been modified from Wodtke (1974) for the isolation of mitochondria in eels. It employs precautions to minimize red blood cell contamination. Eggs, when available, were used in this analysis because of their abundance of mitochondrial DNA. All tissues,

reagents, centrifuge tubes and bottles were kept on ice as much as possible. Ovaries were disrupted, loose eggs were collected and remaining eggs were scraped from the walls of their sacs with a blunt razor blade. From 100-200 grams of tissue were washed several times through two layers of pre-washed cheesecloth with ice-cold 150 mM KCl, 50 mM Tris-HCl, pH 7.4, 2 mM Na_2 EDTA in order to eliminate as much blood as possible. The tissue was then mixed in a Waring blender with MSTE buffer (225 mM mannitol, 75 mM sucrose, 50 mM Tris-Cl pH 7.4, 2 mM Na_2 EDTA) for 20 seconds and aliquots were then homogenized with a Dounce homogenizer to disrupt cells. The blending and homogenizing were done in a refrigerated cold room. Homogenates were brought to 10% (wt. of tissue/vol.) with MSTE and centrifuged 3-4 times (or until no pellet was detected) for 10 min. at 2500 RPM in a Sorvall GSA rotor at 4°C. The supernatant was retained after each spin and precipitates containing fat, nuclei, and other cellular debris were discarded. The debris-free supernatant was then spun at 9500 RPM. for 40 min. The resulting reddish-brown mitochondrial pellet was resuspended in 15 ml of MSTE buffer and layered on top of a preformed sucrose step gradient (10 ml of 1.5 M and 10 ml of 1.0 M sucrose) in TE buffer (10 mM Tris-Cl pH 7.4, 1 mM Na_2 EDTA). This was spun at 25,000 RPM for one hour at 4°C in a Sorvall AH 627 rotor (37 ml buckets) in a Sorvall OTD 65 ultracentrifuge. Purified intact mitochondria were at the interface of the two sucrose solutions. The top layer was suctioned off by vacuum and discarded. The mitochondrial layer at the interface was removed by using "shorty" serological pipets and diluted approximately 4:1 with MSTE buffer and centrifuged for 15 minutes at 13,000 RPM in a Sorvall SS 34 rotor. The resulting mitochondrial pellet was kept for lysis.

Mitochondrial DNA Isolation

The final pellet was resuspended in 3 ml of STE buffer (100 mM NaCl, 50 mM Tris-HCl pH 8.0, 10 mM Na_2 EDTA). Twenty % SDS (BDH) was added to a final concentration of 2% to lyse the organelles, and the lysate was incubated at 37°C for 20 minutes or until clearing was observed. One-sixth volume of saturated CsCl in TE buffer was added to the lysate. The lysate was then stored at 4°C for at least one hour to allow for formation of a Cs dodecyl SO_4 complex containing proteins and membranes. The lysate was then centrifuged at 12,000 RPM in a Sorvall SS-34 rotor for 15 min. Ethidium bromide was added to the resulting clear supernatant to a final concentration of 800 ug/ml and solid CsCl to a buoyant density of 1.57-1.58 g/cm^3. All buoyant densities were determined with the aid of a refractometer. The lysate was then centrifuged at 36,000 RPM for 72 hours at 20°C in a Beckman SW 40 rotor or Sorvall TST 60.4 swinging bucket rotor. Upon completion of the centrifugation run, the mtDNA was present as a very faint band 2-3 millimeters below the brightly flourescent nuclear DNA band, or in many samples was not visible at all. These bands were viewed with short- wave UV light. The mtDNA was recovered via side puncture and its buoyant density was adjusted to 1.35 g/cm^3. This was layered on top of a preformed CsCl step gradient consisting of a bottom layer of 0.7 ml (1.7 g/cm^3) and a top layer of 2.5 ml (1.4 g/cm^3) in a 4.4 ml ultracentrifuge tube. This gradient was spun at 45,000 RPM in a Sorvall TST 60.4 rotor for 3.5 hr. at 21°C. The bottom 1.4 ml volume was recovered via bottom puncture into another ultracentrifuge tube. A layer of 0.8 ml CsCl (1.55 g/cm^3) containing ethidium bromide (200ug/ml) was pipetted on top of each sample, which was then spun for 20 hours at 36,000 RPM in a TST 60.4 rotor. Two bands were evident. The bottom band, removed by side-puncture, consisted of closed, superhelical mtDNA. The top band contained relaxed

mtDNA and sometimes nuclear DNA contamination (Avise and Lansman 1983). During this last centrifugation, cellulose tubes manufactured by Sarstedt were used, because they allowed for better visualization of the two bands than polyallomer tubes. The Beckman cellulose tubes had a tendency to develop stress lines which hindered visualization of the bands and thus were not used. (This difficulty was also noted by Densmore et. al. (1985)).

Ethidium bromide was then removed through 4 to 5 steps of extractions, using butanol containing saturated CsCl. CsCl was extracted through dialysis using the multiwell Bethesda Research Laboratories (BRL) Dialysis Apparatus Model 1200 MD and the appropriate dialysis filters. The dialysis buffer contained 50 mM NaCl,10 mM Tris-Cl, pH 7.5, and 1 mM Na_2 EDTA. The DNA was then precipitated by centrifugation with first 100% and then 70% ethanol, dried, and resuspended in TEN buffer (10 mM Tris-HCl pH 7.4 , 1 mM EDTA, 1 mM NaCl).

Restriction Endonuclease Digestion Procedures

The restriction enzymes Ava II, Dde I, EcoR I, Hinf I, and Rsa I were tested on the samples (see Rationale for Choosing Restriction Endonucleases section for numbers of samples digested with each enzyme). EcoR I is a six-base-, Ava II, Dde I and Hinf I are five base-, and Rsa I is a four-base-cutting restriction endonuclease. The recognition sites from 5' to 3' for each enzyme are as follows with an apostrophe indicating the cut site: Ava II, G'GACC; Dde I, C'TNAG; EcoR I, G'AATTC; Hinf I, G'ANTC and Rsa I, GT'AC.

All digests were performed according to manufacturers' specifications (International Biotechnologies Inc. (IBI), New England Biolabs (NEB), Bethesda Research Laboratories (BRL) or Boehringer Mannheim Biochemicals (BM). Approximately 1-8 units of enzyme were used per digest. Digestion times

varied between 2-4 hours. MtDNA concentrations were not determined, but ranged from between 2 and 10 ng according to visualization next to known concentrations of standards.

The large fragment (Klenow) of DNA polymerase I was used to end-label the 3' termini of the DNA fragments with ^{32}P deoxyribonucleoside triphosphate precursors (Maniatis et al. 1982). Radiolabeled digests were alcohol precipitated, dried, and resuspended in 1 X TBE buffer (8.9 mM Tris, 8.9 mM boric acid, 25 mM Na_2 ETDA, pH 8.3) and stop load (50% glycerol, 0.25% bromphenol blue, 0.25% xylene cyanole).

Separation and visualization of fragments

MtDNA fragments were separated by gel electrophoresis. Large fragments resulting from six-base restriction endonuclease digestions were separated on agarose gels (1%) made in TBE buffer Fragments generated by four-and five base-restriction endonuclease digestions varied sufficiently in size such that their visualization required the use of both agarose (1.4-1.6%) and polyacrylamide gels (6%). Polyacrylamide gels, due to their restrictive porosity, allowed the visualization of the smaller (less than 500 bp) fragments. A 1 kilobase-pair ladder (BRL) and/or lambda DNA digested with Hind III or with Hind III and EcoR I, phi-X 174 digested with Hae III, or the plasmid pBR322 digested with Pst I and Dde I were used as molecular weight markers. All gels were dried under vacuum, either using a homemade drying device or a Hoeffer large-slab gel dryer, Model 1160. Fragments were visualized by autoradiography with (at -80 C) and without intensifying screens (at room temperature) using Kodak XAR 5 film. Mobility of all mtDNA fragments was determined either manually or through use of a sonic digitizer (IBI). Laser densitometry (LKB Ultrascan) was performed to identify some comigrating

bands.

Rationale for Choosing Restriction Endonucleases

Various restriction enzymes were tested on a subset of samples (approximately 50) to determine if they revealed polymorphisms. The restriction enzymes Ava II, Dde I, EcoR I, Hinf I, and Rsa I were then chosen to be tested on all samples because they revealed polymorphisms in this subset.

Because EcoR I, a six-base cutting restriction enzyme, produced only three or four fragments as revealed by agarose gels, it was probable that more polymorphisms would be detected when four or five-base cutting enzymes were used because of the greater number of expected fragments. Thus Ava II, Dde I, Hinf I, and Rsa I were chosen. Both agarose (ag) and polyacrylamide (ac) gels were initially run on samples digested with Ava II, but since there were so few polymorphisms depicted on the polyacrylamide gels, only the agarose gels were continued for all samples. An assumption was made that approximately the same number of fragments was found in each genotype. Since 50 samples digested with Rsa I and run on agarose gels revealed no polymorphisms, agarose gels were discontinued with this enzyme and only the polyacrylamide fragments were scored. Two other restriction enzymes, Ava I, and Bgl I, produced monomorphic fragment profiles when tested on more than 50 samples and were not used further. Taq I and Msp I digestions of over 50 samples revealed polymorphic patterns but due to time constraints and the finite amount of DNA available, the use of these enzymes was discontinued.

Data Analysis

A. MtDNA genotypic frequencies revealed by single enzymes: significance of differences

Digest patterns generated from single enzymes were generally identified by a single letter. Due to the large number of polymorphic fragments which were generated by Dde I and Hinf I digestions, separate letter designations were assigned to Dde I and Hinf I genotypes scored on agarose and polyacrylamide gels, respectively. Additionally, multiple letter designations were assigned to mtDNA genotypes revealed by Dde I and scored on polyacrylamide gels in order to show the commonality of a single DNA fragment (see genotype F) within a spawning system (the Miramichi River). For all enzymes, the A genotype was the most frequent among all samples tested.

Frequencies of each mtDNA genotype determined by single restriction enzymes in each river were determined. For each set of single restriction enzyme genotype frequency data in which there were two or more genotypes, genotypes were pooled, to offset 0 frequencies, and a goodness-of-fit (G test) was used to determine the heterogeneity of the data sets. A program entitled R X C in the statistical package (BIOM) (Sokal and Rohlf 1981) was used for "unplanned tests of the homogeneity of replicates tested for goodness of fit". The replicates in this case were the genotypic frequencies in each river. The program requires that a critical G be entered (for an alpha of 0.05) which is comparable to the critical X^2 as determined by a X^2 table (Sokal and Rohlf 1969). The frequency data was compared with this critical G.

First, the genotypic frequencies for shad in all nine rivers were tested to see if there was overall heterogeneity. The G-test was then used to test heterogeneity between all possible combinations of mtDNA genotypic frequencies as revealed by single restriction enzymes in shad from individual

rivers. The program did this by deleting shad mtDNA genotypic frequencies from one river at a time, until all homogeneous (non-significant) subsets were found.

Another test, the Kolmogorov-Smirnov (K-S) two-sample test, was used which is recommended for sample sizes of less than 50 and small sample sizes in individual cells (Sokal and Rohlf 1981). This is a non-parametric statistic that makes no assumptions about the underlying distribution of the data. This test is based on the absolute value of the differences between the relative cumulative frequency distributions of two collections. The maximum difference when comparing cumulative frequencies of the genotypes in the two populations is determined to be the critical value Z, and is compared in a table for significance (Sokal and Rohlf, 1981). An SPSS program was used to generate K-S values.

Scatter diagrams of the latitude of the collection site vs. frequency of shad mtDNA genotype in the appropriate river were plotted using the program Cricket Graph (Cricket Software). The single restriction enzyme genotype data that was used for these analyses were: Ava II, Dde I agarose, Hinf I agarose and polyacrylamide, and Rsa I. These genotypic frequencies were chosen because of their apparent correlation with latitude when rough plots were examined by eye alone.

After the scatter plots were made, the linear regression function was chosen, which plotted the best straight line through the points. The linear regression equation and the r^2 value, which is the coefficient of determination, which ranges from 0 to 1, also are shown on the graphs. The r^2 value (X 100) measures the proportion of the Y sums of squares (sum $(Y - Y)^2$) that is "explained" by the regression formula (Leaverton, 1986).

The linear regression lines and coefficients of determination were appropriate only under the assumption that the data were normally distributed,

and that there was a linear relationship between latitude and genotypic frequency. As that information was not known, the Kendall's rank-coefficient of correlation was also calculated for each comparison. This is a non-parametric test.

The following procedure from Wardlaw (1985) illustrates how to obtain this coefficient. Rivers were listed from lowest to highest latitude, along with the corresponding genotype frequencies. The value P, the number of larger frequency values below the one listed was entered in the space adjacent to the corresponding frequency. The value Q, the number of smaller frequency values below the one listed was entered in the column to the right of P. The column P-Q was obtained by subtracting the two values, which were then summed to give the value S. The Kendall's rank-correlation coefficient K was then calculated by:

$K = 2S/N(N-1)$

where N equals the number of river-genotypic frequency values.

The statistical significance of K for an N less than 10 (as was in this case) was determined by looking up S in Table A10, entitled "Probabilities (per cent) for testing the significance of S, obtained in calculating Kendall's coefficient of rank correlations". (Campbell 1967).

B. Composite mtDNA genotypes: estimates of diversity

Each fish was assigned a multi-letter composite genotype based on the combination of results of all individual restriction enzyme digests. An estimate of restriction pattern diversity within spawning systems was calculated (Nei and Tajima 1981). The formula for this calculation is as follows:

$h = N(1 - x_i^2)/N - 1$

where h is the restriction pattern diversity, N is the total number of samples per river, and x is the frequency of the ith composite mtDNA genotype.

Because of the large number of unique composite genotypes, phylogenetic relationships among every fish were not determined. Thus, composite genotypes found in more than one river ("shared" composite mtDNA genotypes), which represented a workable subset of all of the composites, were examined to see if they were concordant with geographic partitioning. For example, were these genotypes shared only among adjacent rivers or among rivers of a particular geographic region?

Before the relationships were ascertained, it was necessary to quantify how much each genotype diverged from the others. A genetic divergence value (p) expressed as a percentage value indicates the proportion of base pairs in a mtDNA molecule that have changed over time. In order to calculate p, the F statistic (fraction of shared fragments) (Upholt 1977) was calculated. The F statistic quantifies the extent of mtDNA restriction fragment similarity between individuals having different genotypes via the relationship:

$F = 2 N_{XY} / (N_X + N_Y)$

where N_X and N_Y are the number of fragments scored in genotypes X and Y, and N_{XY} is the number of fragments shared between the two genotypes.

These F values were converted to estimates of base sequence divergence (p) by the following relationship:

$p = -(\ln F)/r$

where r equals the numbers of bp at the recognition site of the particular enzyme being used (Nei and Li 1979). The conversions were made using tables compiled by Nei (obtained from J. C. Avise) that had separate conversions for each F-p value based on how many bp each restriction enzyme recognized. The p value corresponds to the percentage of the mtDNA molecule that has changed between the two genotypes being compared.

In order to convert F values to p (divergence) values the assumption was

made that the fragments for each comparison were homologous and unexplained polymorphisms were attributable to site variations or one mutation event. Templeton (1983) suggested convergent evolution of restriction sites (such as a gain-loss-gain of the same restriction site) can be a problem in this interpretation. In addition, restriction sites are not always randomly distributed throughout the mtDNA molecule, an assumption of this approach (Adams and Rothman 1982). Finally, length mutations are not accounted for in the Nei and Li (1979) method. All of these problems can introduce some unknown error into the data.

For these reasons, site data are preferable to fragment data, because the latter may actually overestimate divergence. For example, in the simplest case of comparing a mtDNA molecule with one restriction site to a second mtDNA molecule with the same restriction site plus one more, the F value would be 0.66 when comparing site data but zero when comparing fragment data. The larger fragment of the first mtDNA molecule would be compared to the two smaller fragments of the second mtDNA molecule; thus no fragments would be shared among the two genotypes.

However, in order to use site data, a map of the restriction sites on the mtDNA molecule must be obtained. Because of the limited amount of mtDNA available, time constraints, and the numerous fragments revealed by the enzymes used in this study, it was impossible to map each mtDNA molecule and do the appropriate digestions in order to determine if polymorphisms are present. Similar problems were encountered by Zink and Avise (1990). Southern-blotting using homologous probes, a technique not employed in this analysis, can be used to identify homologous fragments. However, the materials needed for this type of analysis were not readily available at the start of this study. Monomorphic restriction enzyme profiles (as determined by a

subset of 50 fish) were also not used in this analysis, so divergence values may actually be overestimated.

If one assumes a clock of 2% base sequence change per million years (Brown et al.1979) then one can estimate the time it took for two genotypes to diverge. However, Dowling and Brown (1989) stipulate that different parts of a mtDNA molecule can possibly undergo different rates of base substitutions.

C. Phylogenetic analysis

Both distance and parsimony methods (reviewed in Felsenstein (1988) were used to construct phylogenies between and among shared genotypes. P values were used to construct a phenogram that related the shared composite genotypes using the Fitch-Margoliash distance analysis component of the Felsenstein PHYLIP program (Fitch and Margoliash 1967). A distance method groups genotypes based on overall similarities (in this case, similarity of proportions of shared fragments). The algorithm successively adds each p value entered, and the best tree is considered to be that which minimizes the sum of squares of the differences between the distances in the tree and the distances in the input p values.

In contrast to the distance approach, a parsimony analysis using the HENNIG86 program, version 1.5 (available from J. S. Farris, Dept. of Ecology and Evolution, State University of New York, Stony Brook, NY 11794) that related the shared composite genotypes was also determined. This was based on the presence (1)/absence (0) of fragments in each genotype among individual comparisons. Only the fragments that varied across genotypes (considered to be shared derived characters or synapomorphies) were used in this analysis. Rare fragments (those that occurred only once and considered to be autapomorphies) and fragments shared by all genotypes were deemed

uninformative. The model for using this analysis with mtDNA composite genotypes as taxa and fragments (presence or absence of as characters) was found in Riddle and Honeycutt (1990). The parsimony analyses are also sensitive to problems of interpretation when fragment data rather than site data are used. As in the phenograms, fragments with similar mobilities were assumed to be homologous in the initial coding. However, it is likely that fragments with high levels of homoplasy are not truly homologous (Patterson 1989). It is probably a larger problem when dealing with fragments that do not show identical mobility because of small inserts but do share identical sites at their termini.

A parsimony analysis attempts to minimize the overall tree length. An assumption made with Wagner parsimony analysis is that it is equally as likely for a restriction site to be lost as it is to be gained, which may be incorrect, as it is more probable that a restriction site will be lost rather than gained. A site will be lost if only one base pair is changed whereas the proper 4-, 5- or 6-base pair sequence must be gained in order for a new site to be created. Dollo parsimony is often used to correct for this problem, but cannot be used here because it requires site data.

Characters that are more reliable, i. e. show the least reversals or convergences, produce a tree with a higher overall consistency index, defined as:

CI_j = sum r_j/sum l_j

where r is the range of character i which is the minimum number of steps that the character would need to be placed on a given tree and l is the length (number of steps) required by the given tree (Kluge and Farris 1969). The overall consistency index was calculated for each tree.

A retention index was also calculated for each tree. A retention index is a

measure of how often a character is synapomorphic in a suite of composite mtDNA genotypes. It is defined as:

$RI_i = (h_i-s_i)/(h_i-1)$

where h_i is the maximum number of steps for character i, and s is the total number of steps in the tree, with 1 being the minimum number of steps (Farris 1989). A unique character would have an RI of 0 and a character shared by all genotypes would have an RI of 1.

Successive approximations (Farris 1969) were used to eliminate some of the equally parsimonious trees. In this method, the character weights are calculated as the product of the individual consistency and retention indices for each character and scaled between 0 and 10, and the trees are recalculated. This procedure is repeated until the number of equally parsimonious trees stabilizes. This successive approximations method has been shown to be relatively insensitive to the starting tree (Farris and Fitch, unpublished data). There is no algorithm for bootstrapping trees obtained with successive approximations.

MtDNA from one hickory shad (Alosa mediocris) from the Santee-Cooper system was isolated and digested with restriction endonucleases, its fragment patterns were compared to those of the American shad, and appropriate F and p values were derived with the same assumptions as listed above. This taxon served as an outgroup in order to root both the phenograms and cladograms.

RESULTS

Number of base pairs and types of fragments scored

The five informative restriction enzymes used in this survey produced 125 restriction fragments for the most common genotype (A) and varying numbers of fragments for the other genotypes. This was representative of 603 nucleotides, or 3.3% of the approximately 18,300 base pair shad mtDNA molecule (size reported by Bentzen et al. 1988). These were not all the fragments generated by these enzymes. Mitochondrial DNA fragments less than 60 bp were not scored because of poor resolution on the gels and a lack of mtDNA molecular weight standard fragments in that size range. In addition, comigrating bands were frequently generated as indicated by darker or wider bands on the gels, but these were often scored as single fragments. Fragments produced by enzymes that revealed monomorphic patterns were also not considered by this analysis. The mtDNA fragment number and sizes for each genotype uncovered by each restriction enzyme are listed in Table 2.

Types of polymorphisms

Two different types of mtDNA polymorphisms were scored. One form of variation was a definitive restriction site polymorphism, as revealed by EcoR I, in which the number of fragments varied among individuals and in which the sum of the molecular size of two smaller fragments was equal to that of a larger polymorphic fragment. The second class of variants, as generated by Ava II, Dde I, Hinf I and Rsa I, was characterized by differences in the electrophoretic mobility of individual fragments. These were most frequently detected on polyacrylamide gels.

Fourteen mtDNA genotypes were revealed with Ava II. Polymorphisms were primarily observed within two molecular size ranges; 1200-2000 bp (Figure 2) and 250-800 bp (Figure 3). In both cases, differences were detected in the electrophoretic mobility of individual fragments and the absence of single fragments. For example, Genotypes A and B differed in the electrophoretic mobility of an approximately 1400 bp fragment. In addition, Genotype C was missing a 1200 bp fragment that was present in all other genotypes. However, this genotype exhibited a 1510 bp fragment not evident in the other genotypes. Differences were also scored among smaller mtDNA fragments whose molecular size ranged between 250 bp and 350 bp, and which were separated on 6% polyacrylamide gels. For example, Genotype C revealed comigrating bands of 325 bp, while Genotypes A, B, D, and E exhibited 350 bp fragments. Genotype D was missing a 280 bp fragment present in all other genotypes.

Six Dde I digest patterns were differentiated on agarose gels. Genotype A was characterized by the presence of a 540 bp fragment not seen in the B genotype. Instead, Genotype B displayed a 1015 bp fragment (Figure 4). Genotype C was characterized by a 1430 bp fragment, and was missing two fragments, 610 bp and 595 bp, that were present in the other two genotypes. This could be indicative of a site polymorphism.

Eighteen different Dde I genotypes were observed on 6% polyacrylamide gels. Based on molecular sizing, these were different mtDNA fragments than were scored on agarose gels. Variation was seen in the electrophoretic mobility of fragments which ranged from about 250 to 530 bp (Figure 5). The Genotype F pattern was missing a 245 bp fragment that was present in all other genotypes. (The Genotypes BGF and EF were also missing this 245 bp fragment, but they also had fragments in common with the other lettered genotypes.) This could have been indicative of heteroplasmy (more than one

size-class of mtDNA in one individual) for electrophoretic mobility variants. This was evident in twelve individuals (several of these samples were run on gels more than once to confirm heteroplasmy).

EcoR I revealed two genotypes resulting from a restriction site polymorphism. Genotype A was characterized by three fragments on agarose gels; 9970 bp, 7315 bp, and 1370 bp. Genotype B was characterized by four fragments, the largest mtDNA fragment (9970 bp) seen in Genotype A was cleaved into two smaller mtDNA fragments of approximately 5100 bp and 4220 bp.

Agarose gel electrophoresis of Hinf I digested mtDNA revealed four genotypes, two of which were found in two Santee-Cooper fish only. Genotype A lacked a 1065 bp piece that Genotype B had, and Genotype B was missing a 1040 bp fragment that was present in Genotype A (Figure 6). These genotypes were also inherited independently of the Hinf I genotypes described below that were depicted on polyacrylamide gels. Thus, different combinations of Hinf I-agarose-Hinf I-polyacrylamide genotypes were evident.

Separation of Hinf I digests on polyacrylamide gels revealed two genotypes (Figure 7). Genotype A lacked a 223 bp fragment that was found in Genotype B, while Genotype B lacked a 185 bp fragment that was present in Genotype A. This accounted for a 50 bp difference in size for the two molecules as depicted on polyacrylamide gels.

Rsa I digests produced twelve different genotypes on polyacrylamide gels. Variation was seen in fragments which ranged in size from 300 bp to 365 bp. Variation in the electrophoretic mobility of these small fragments could be attributed to differences of 10 bp to 15 bp in the size of the polymorphic fragments (Figure 8). Some genotypes lacked fragments present in the other genotypes. An example was Genotype D, which had no fragments sized from

325-330 bp. Heteroplasmy for electrophoretic mobility variants was evident in fourteen individuals from four rivers (several of these samples were run on gels more than once to confirm heteroplasmy). These were scored as multiple letter genotypes in Table 3. Three different mtDNA fragments were detected as the heteroplasmic mtDNA fragment in these individuals. These same individuals were not heteroplasmic using other restriction enzymes.

Frequencies of polymorphisms

Frequencies of individual restriction enzyme digest pattern (Ava II, Dde I (agarose and polyacrylamide), EcoR I, Hinf I (agarose and polyacrylamide), and Rsa I (polyacrylamide) are listed in Table 3 a. Frequencies of individual restriction enzyme heteroplasmic polymorphisms are listed in Table 3 b.

The goodness-of-fit tests revealed overall heterogeneity in shad genotypic frequencies among rivers with the enzymes Ava II, Dde I (on agarose gels), Hinf I (on acrylamide gels) and Rsa I. Non-significant subsets are depicted in Figure 9.

The results of the K-S tests which list significance of differences in the frequencies of these genotypes for between-river comparisons are given in Tables 4. The K-S values for two separate enzymes were listed above and below the diagonal respectively.

Ava II genotypes A and B were detected in 83% of all samples, with Genotype L being the next most frequent at 8%. The goodness-of-fit test revealed eight homogeneous subsets. Significant heterogeneity in genotypic frequencies was detected among comparisons between data from the St. Johns, the Altamaha, the Chowan/Albermarle Sound. the Delaware, and the Hudson alone and the other rivers collectively. Only when these rivers were removed individually from the analysis, were homogeneous subsets of the other

riverine genotypic frequencies observed. The removal of the the Santee-Cooper Rivers and one of three other rivers (the Miramichi, the Connecticut or the Rappahannock) revealed homogeneous subsets with the remaining genotypic frequencies. Using the K-S test, there were six significant differences in frequencies of shad mtDNA genotypes, five between pair-wise comparisons of mtDNA genotypic frequencies of shad from the Chowan River/Albermarle Sound with those from other rivers.

The frequency of Dde I agarose genotype B decreased from south to north and was absent from the Miramichi River collection. Genotype C was found in four fish which comprised 27% of the Miramichi River samples and in one fish from the Chowan River/Albermarle Sound and two from the Delaware River. Seven homogeneous subsets of shad mtDNA genotypic frequencies were revealed with the goodness-of-fit test, due to the removal of more than one set of riverine shad mtDNA genotypic frequencies at a time. There was one southern contiguous subset of shad mtDNA genotypic frequencies which included data from the St. Johns through the Rappahannock rivers. There were nine significant differences between pair-wise comparisons of shad mtDNA genotypic frequencies depicted with the K-S test, four between the St. Johns River shad mtDNA genotypic frequencies and other riverine shad mtDNA genotypic frequencies and five between the Chowan River/Albermarle Sound shad mtDNA genotypic frequencies and other riverine shad mtDNA genotypic frequencies.

Dde I polyacrylamide Genotype A ranged in frequency from 62% to 84% in the American rivers, and then showed a dramatic dip to 31% in the Miramichi River. Genotype C was found in all rivers except the St. Johns. Eight of 13 (62%) Miramichi River fish contained some form of the F genotype, which was not found in any other river. The goodness-of-fit test could not be used at all on

the Dde I polyacrylamide data because of the many cells with expected frequencies of zero, whereas the K-S test was able to differentiate the Miramichi River shad mtDNA genotypic frequencies from those of all other fish.

EcoR I Genotype A was found in fish from all rivers, but Genotype B was found only in nine fish, eight from the St. Johns River and one from the Hudson River. There were no significant difference detected in the between-river shad mtDNA genotypic frequencies comparisons with the K-S test.

Genotype A as revealed on Hinf I agarose gels was the most frequent in shad from all rivers except the Delaware, where it was present in 45% of the fish. Genotypes C and D were found in only one fish each--both from the Santee-Cooper rivers. The goodness-of-fit test with the Hinf I agarose data revealed one homogeneous subset comprised of all riverine genotypic frequencies. Frequencies of Hinf I agarose genotypes revealed only one significant difference when tested with K-S, that between fish from the Delaware River and the Chowan River/Albermarle Sound.

Genotype A was the most abundant Hinf I polyacrylamide genotype at a frequency of 87%. Genotype B was discontinuously distributed along a geographic continuum and appeared in shad from the Santee-Cooper rivers, Chowan River/Albermarle Sound and the Hudson-Delaware rivers and was either absent or present in very low frequencies in the St. Johns, Altamaha, Rappahannock, Connecticut and the Miramichi rivers. The K-S test on comparisons with Hinf I polyacrylamide data revealed four significant differences, all between the Chowan River/Albermarle Sound shad mtDNA genotypic frequencies and those of other rivers.

With the Rsa I data, Genotype A was the most abundant in Miramichi River shad (73%), while only 4 (12%) of the St. Johns River samples exhibited this genotype. However, the frequency of this genotype showed a wide

variation among all the other rivers. Genotype B was found in the highest frequency in the Santee-Cooper rivers and Chowan River/Albermarle Sound regions (61% and 65%, respectively). Genotype C was found with the greatest frequency in fish at the southernmost end of the shad's range--50% in the St. Johns River and 25% in the Altamaha River. It was also found in shad from all other rivers except the Miramichi. Genotype E was found only in 3 fish, all from the Miramichi River(20%).

Twenty homogeneous subsets were found with the goodness-of fit tests on Rsa I shad mtDNA genotypic frequency data. None of these subsets was homogeneous when only one set of riverine shad mtDNA genotypic frequencies at a time was excluded from the comparisons. The Santee-Cooper rivers, Chowan River/Albermarle Sound and the Delaware River shad mtDNA genotypic frequencies were significantly different from at least eight (the least number) subsets of data, whereas the Miramichi River data set was significantly different from at least fourteen subsets of shad mtDNA genotypic frequencies. The St. Johns River shad mtDNA genotypic frequencies were not significant from all but two sets of riverine shad mtDNA genotypic frequencies in any one comparison.

The K-S test differentiated the St. Johns River shad mtDNA genotypic frequencies from those of all other rivers in the pair-wise comparisons.

Correlation of genotype with latitude

Linear regression analysis was used to compare individual enzyme genotypic frequencies with the latitude of each spawning river. The highest r^2 value was 0.532 for the Dde I agarose data for the B genotype (Figure 10). The lowest value was 0.005 for the Hinf I acrylamide data. The probability that the Dde I genotype B data did not correlate with latitude was 0.024 ($p < 0.05$), the

only significant data for the correlation analysis. Tables 5 and 6 summarize the correlation analysis.

Frequencies of composite genotypes

Composite genotypes were compiled for shad in each of the nine spawning rivers (Table 7). We were unable to resolve digest patterns for all the enzymes on all samples and thus composite genotypes (98) were reported for only 161 of the samples. Table 7 also lists the total sample number for each river, the number of shad mtDNA genotypes in each river, and the number of genotypes that are unique to each river. The unique genotypes can be assigned to more than one sample. The total number of unique genotypes was ninety (56%). The number of rare genotypes (those that occurred only once) was 78 (48%). The St. Johns River revealed the largest percentage of within-river and unique composite genotypes. A smaller percentage of fish from more northerly rivers displayed unique genotypes (35% to 60%). The Delaware and the Hudson rivers samples displayed the lowest percentages of unique genotypes. No single composite genotype was seen in shad from all nine rivers and only 15 of the 98 composite genotypes were observed in shad in more than one river. The frequency of individual composite genotypes never exceeded 15% in shad from any one river.

Restriction pattern diversity

The estimate of restriction pattern diversity (Nei and Tajima 1981) based on the number of samples analyzed and the number and frequency of composite genotypes in individual spawning systems is listed in decreasing order: St. Johns River (0.986), Chowan River/Albermarle Sound (0.985), Miramichi River (0.978), Delaware River (0.971), Santee-Cooper rivers (0.964),

Hudson River (0.964), Connecticut River (0.964), Rappahannock River (0.954), and Altamaha River (0.952).

Phylogenetic analysis

Because of the large number of unique composite genotypes, only those composites that were shared among rivers (Table 8) were used for phylogenetic analyses in an attempt to uncover relationships that could be concordant with geography. Figure 11 depicts the shared composite genotypes geographically.

Shared fragment (F) and divergence (p) values

The fraction of shared fragments (F values) (Table 9 above diagonal) ranged from 0.94 to 0.993 for the fifteen shad composite mtDNA genotypes shared among rivers. Divergence (p) values (Table 9 below diagonal) ranged from 0.00057-0.00422.

Distance phenogram

A Fitch-Margoliash phenogram was calculated from the distance (p) values obtained by comparisons between the 15 shared composite genotypes. An unrooted phenogram (Figure 12 A) and a phenogram (Figure 12 B) using the hickory shad as its root were constructed with the PHYLIP program. A percent-divergence scale was indicated on each figure.

Parsimony cladogram

The synapomorphic fragments for each individual-enzyme genotype (Table 10) were coded by one/zero (presence/absence of each fragment) (Table 11) and combined for each of the 15 shared composite genotypes (Table 12). Three unrooted equally parsimonious trees (Figure 13) and one tree

rooted with the hickory shad (Figure 14) were calculated based on successive approximations of the characters using the HENNIG86 program.

Comparisons of divergence values of American shad to hickory shad

F values ranged from 0.133 to 0.461 and p values ranged from 0.055 to 0.152 between the two Alosa species (A. sapidissima and A. mediocris). These results were based on one comparison only between the one A. mediocris genotype and one A. sapidissima genotype on a single gel for each enzyme. Shared fragments were scored based on comigrating fragments for the two samples (See Table 13) ; however, since homology of the two fragments for each comparison is unknown, divergence values could actually be much higher between the two species.

DISCUSSION

Identification of stocks

The hypothesis that geographically separate shad spawning rivers contain discrete populations was tested. Implicit to this study was the assumption that spawning fish collected at spawning sites could be considered natal to that system. Mark-recapture studies have shown that American shad are relatively faithful homers (about 97%) (Hollis 1948; Mansueti and Kolb 1953; Williams and Daborn 1984; Dodson and Leggett 1973; Melvin et al. 1986). Meristic data have suggested that shad not only home to their natal rivers but also to tributaries within the St. John River, New Brunswick. (Carscadden and Leggett 1975). However, U. S. Fish Commission reports (1871-1886) show that extensive shad stocking programs were conducted along the U. S. eastern coast. In come cases, eggs, juveniles, and adult fish were transplanted to non-natal rivers. This could have been a source of "artificial infidelity" that could complicate stock identification. Even low levels of infidelity may be sufficient to prevent the establishment of river- specific genotypic frequencies (Allendorf et al. 1983).

American shad stock discrimination encompasses two related but separate questions that require differing degrees of resolution to address. First, can separate stocks be identified? The answer requires the demonstration of significant differences in the frequency distributions of a single or composite polymorphic markers. An affirmative finding indicates that these discrete populations or "stocks" could then be managed as separate units. The second task requires greater resolution. Can a reasonable percentage of fish from these defined stocks be uniquely assigned an origin in a mixed ocean fishery?

To do this, one must estimate the relative contribution of the defined stocks to the mixed ocean fishery. The identification of ancestry of individuals or contingents of fish in a mixed offshore fishery may require the demonstration of individual stock specific markers that are minimally at a moderate frequency within the previously defined stocks or the detection of a variety of polymorphic markers, each of which differs in frequency among the stocks tested. In order to decide with a reasonable probability that an unknown sample came from a particular stock, the stocks themselves must be able to be differentiated 95% of the time. Even if stocks are identifiable 70 % of the time, errors are compounded in stock assignment (T. L. Ong and S. Saila, personal communication). If an excessively high level of variability is detected within spawning systems, individual stocks can still be defined, but the identification of ancestry of individual fish in a mixed fishery becomes more complex.

To answer the first question, tests of significance of single-enzyme genotypic frequency distributions between rivers were conducted using the goodness-of-fit (X^2) and K-S tests. The goodness-of-fit analyses were conducted on the Ava II, Dde I agarose, Hinf I agarose, Hinf I polyacrylamide and Rsa I data sets. Pooling of some genotypes was conducted with Ava II, Dde I agarose Hinf I agarose and Rsa I data to offset zero cell values. The goodness-of-fit tests were not done on Dde I acrylamide data because there were too many genotypes to justify pooling; the K-S test was used instead. The goodness-of-fit tests were not carried out on EcoR I and Hinf I polyacrylamide data as there were many zero frequency values.

Because all "not A" genotypes were pooled with the Ava II, Dde I agarose and Hinf I genotypes and A, B and "not A or B" with the Rsa I data some discriminating power was lost with the goodness-of-fit analysis. Pooling masked heterogeneity, rare genotypes, and the rather high frequency of a few

genotypes (see Genotype L for the Ava II analysis of the Chowan River/Albermarle Sound). Some power was gained with the use of the K-S test because genotypic frequencies were not pooled; however, the latter test is less sensitive.

The advantage of the goodness-of-fit tests was that all frequency data from every river could be tested for overall heterogeneity simultaneously. If heterogeneity was present, the source could be determined by excluding one river at a time from the analysis until all homogeneous subsets could be determined. In this manner, the source/s of significance could be determined with a minimal amount of error as a large group of data could be analyzed at once. With pair-wise comparisons (by river) done with the K-S tests, there is a higher chance for error for each comparison (Sokal and Rohlf 1981).

The goodness-of-fit tests revealed heterogeneity among riverine mtDNA genotypic frequency comparisons with Ava II, Dde I (agarose) and Rsa I, as evidenced by plural homogeneous subsets. Hinf I agarose data did not reveal a heterogeneous G as shown by the one homogeneous subset of all data.

When the St. Johns, Altamaha, Chowan River/Albermarle Sound, Delaware or Hudson rivers Ava II data sets alone were eliminated from the unplanned tests of replicates, homogeneous subsets of the remaining rivers resulted. This indicates that each of these riverine data sets alone is significantly different from the other riverine data sets combined. Exclusion from three riverine data sets of a combination of the Santee-Cooper rivers data set with those of three other distant rivers--the Rappahannock, Connecticut and Miramichi (in separate comparisons) resulted in homogeneous subsets.

Seven homogeneous subsets were revealed with Dde I agarose data; none with the removal of one riverine data set alone. This indicates that combinations of heterogeneous groups needed to be removed from the

analysis in order to obtain non-significant subsets. As the Connecticut River data set was absent from six of the seven non-significant groupings, the Connecticut River data set was the most significantly different from the other data sets. Frequencies of shad mtDNA Dde I genotypes from the Connecticut River could be distinguished from those of the Chowan River/Albermarle Sound because they did not form a homogeneous subset even with the removal of combinations of rivers from the analysis. The only geographically-clumped homogeneous subset of frequencies of mtDNA genotypes as revealed by the goodness-of-fit analysis was found in shad samples from the St. Johns through the Rappahannock rivers when the frequencies of mtDNA genotypes from the northern rivers were removed from the analysis.

Rsa I data revealed the most heterogeneity among rivers with twenty homogeneous subsets. The Miramichi River data set was absent from the greatest number of non-significant subsets (14), followed by the Connecticut River with 13 and the St. Johns River with 12. The least number of exclusions from non-significant subsets was 8, and included the Santee-Cooper rivers, Chowan River/Albermarle Sound and Delaware River data subsets. The St. Johns data never formed homogeneous subsets with more than one river at a time, indicating that it was significantly different from large groupings of riverine data.

Overall, the unplanned tests of replicates using the goodness-of-fit analysis revealed that there was heterogeneity among genotypic frequencies among rivers. Shad from the St. Johns, Altamaha, Chowan River/Albermarle Sound, Delaware and Hudson systems separately could be distinguished from assemblages of all other rivers based on these tests. Frequencies of shad mtDNA Dde I genotypes from the Connecticut River could be distinguished from those of the Chowan River/Albermarle Sound. There was a geographic

continuum of homogeneous frequencies of mtDNA Dde I genotypes found in shad samples ranging from the St. Johns to the Rappahannock rivers when the frequencies of mtDNA genotypes from shad from the northern rivers were removed. Rsa I was the enzyme that revealed the most heterogeneity, with all rivers each excluded from at least eight homogeneous groupings.

The K-S pair-wise test of the significance of frequencies of shad mtDNA genotypes as revealed by single restriction enzymes was used to determine the following differentiations between riverine populations of shad. Shad from the St. Johns were differentiated from all other rivers with Rsa I polyacrylamide data. Shad from the Altamaha River were unable to be differentiated from other rivers using this test. Fish from the Santee-Cooper rivers were differentiated from those in the Chowan, Rappahannock, Hudson and Miramichi rivers. The Chowan River/Albermarle Sound appeared to be a unique stock that was differentiated from all rivers except the Altamaha River by its frequency distributions using a combination of Ava II, Dde I agarose and Rsa I genotypes. The Rappahannock, Delaware, Hudson and Connecticut rivers (in the mid-range of this study) were only able to separately be differentiated from one or two rivers in the pair-wise K-S comparisons of frequencies of shad mtDNA genotypes. Shad from the Delaware and the Hudson rivers could not be distinguished from each other with this test. The mtDNA genotypic frequencies revealed on Dde I polyacrylamide gels differentiated the Miramichi River shad stock from all others with the exception of shad from the Altamaha River.

Fixed single-enzyme or composite mtDNA genotypes were not detected in any shad spawning system. Two single-enzyme generated mtDNA genotypes were restricted to a single river (except for one sample) and were found in a moderate frequency within that system. The EcoR I B genotype was seen in only 11 of 257 fish tested, 10 from the St. Johns River samples (27%) and 1

from the Hudson River(3%). A combination of F genotypes (Dde I polyacrylamide data) was seen in 8 out of 209 fish tested, all from the Miramichi River (62%). These two polymorphisms could potentially be used to identify St. Johns River and Miramichi River fish from others in a mixed sample.

Bentzen et al. (1989) reported significant heterogeneity among frequencies of four shad mtDNA genotypes among 14 Northwestern Atlantic rivers. Two of these genotypes were single-enzyme (Hae II and EcoR V) site variants and one was a heteroplasmic length variant revealed with Sal I and Kpn I. The fourth genotype was a composite mtDNA genotype composed of mtDNA genotypes produced by four restriction enzymes. Three shad mtDNA genotypes, the two single-enzyme site variants and the length variant were highest in frequencies in particular geographic areas. Frequencies of the Hae II B shad mtDNA genotype were highest in St. Lawrence and Richelieu rivers, and varied significantly (in this area) from those in other rivers. The EcoR V B genotypes occurred in rivers in the mid-Atlantic range of the shad only (the Wacamaw River, South Carolina to the St. John River, New Brunswick). The highest frequencies for this genotype were from the Hudson, Connecticut and Annapolis rivers, where spawning times were also coincidental. A heteroplasmic length variant revealed with Sal I and Kpn I was found in shad from the Annapolis to the St. Johns rivers, Florida, and occurred with the greatest frequency (50%) in the St. Johns River. Thus Bentzen et al. (1989) was able to distinguish shad populations from two (and possibly three) broad geographic regions (the St. Lawrence-Richelieu rivers (based on frequencies of the Hae II B shad mtDNA genotype, the Wacamaw to St. John (New Brunswick) rivers (based on frequencies of the EcoR V B shad mtDNA genotype, and the Wacamaw to the St. Johns (Florida) rivers (based on frequencies of the heteroplasmic length variant revealed with Sal I and Kpn I).

A large number of shad composite mtDNA genotypes (98 out of 161 fish mtDNA's sampled) was detected with the restriction enzymes Ava II, Dde I, EcoR I, Hinf I, and Rsa I in my study. Even though a large proportion of these genotypes (56%) were found only in a particular river, 46% were rare (occurred only once), which minimizes the value of composite genotypes as stock identifiers. However, one of the fifteen composite genotypes shared in more than one river, AAAAAAA, was exhibited in 11 out of 12 fish from rivers inclusive of and north of the Delaware. Bentzen et al's. (1989) results were similar to mine in that of 52 shad tested with nine restriction enzymes, 8 out of 13 composite mtDNA genotypes observed were rare and 3 were shared among rivers, none of which showed geographic clumping. It has been postulated that spawning populations or "stocks" of shad could represent geographic regions rather than specific rivers. Williams (1985) delineated shad based on otolith analysis, Melvin (1984) characterized separate stocks of shad based on multiple meristic and morphometric characters, and Dadswell et al. (1987) found distinct oceanic aggregates of shad within the overall mixed Atlantic coastal fishery. These researchers delineated shad into mainly three distinct geographic spawning groups (Canadian, mid-U. S., and southern U. S. rivers). The data from Bentzen et al. (1989) and from this treatise has fine-tuned these clusters into even more distinct populations. A combination of significant differences in genotypic frequency distributions between stocks and within-stock unique mtDNA markers for the two studies are summarized as follows: shad in the the St. Johns River could be distinguished from the combination of shad found in all other rivers with significant differences of frequencies of shad mtDNA Ava II genotypes (goodness-of-fit), and from shad in each river with significant differences in frequencies of shad Rsa I mtDNA genotypes (K-S tests), and from shad in all other rivers with Kpn I and Sal I length variants (X^2,

Bentzen 1989); shad in the the Altamaha River could be distinguished from the combination of shad found in all other rivers with significant differences of frequencies of shad mtDNA Ava II genotypes (goodness-of-fit) and from shad from the Miramichi and St Johns rivers by significant differences of frequencies of shad mtDNA Dde I polyacrylamide and Rsa I genotypes, respectively (K-S test); shad from the Santee-Cooper rivers could be distinguished from those of the St. Johns River and Chowan River/Albermarle Sound by significant differences of frequencies of shad mtDNA Dde I agarose genotypes (K-S test), shad from the Chowan River/Albermarle Sound could be differentiated from the combination of shad found in all other rivers and from the Connecticut River alone with significant differences of frequencies of shad mtDNA Ava II genotypes (goodness-of-fit) and from shad from all rivers except the Altamaha in significant differences of pair-wise comparisons of the frequencies of shad mtDNA Ava II, Dde I agarose, Dde I polyacrylamide, Hinf I polyacrylamide and Rsa I genotypes (K-S test), shad from the Rappahannock could be distinguished from the shad from the Chowan River/Albermarle Sound by significant differences of frequencies of shad mtDNA Hinf I polyacrylamide genotypes and from shad from the St. Johns, Santee-Cooper rivers, and Chowan River/Albermarle Sound by significant differences of frequencies of shad mtDNA Rsa I genotypes (K-S test); shad from the Delaware River could be distinguished from the combination of shad found in all other rivers with significant differences of frequencies of shad mtDNA Ava II genotypes (goodness-of-fit) and from shad from the St. Johns River, the Chowan River/Albermarle Sound and the Miramichi River by significant differences of frequencies of shad mtDNA Rsa I, Ava II and Hinf I agarose, and Dde I polyacrylamide genotypes respectively (K-S test); shad from the Hudson River could be distinguished from the combination of shad found in all other rivers by

significant differences of frequencies of shad mtDNA Ava II genotypes (goodness-of-fit) and from shad from the St. Johns and the Santee-Cooper Rivers and shad from the Chowan River/Albermarle Sound by significant differences of frequencies of shad mtDNA Rsa I, and Ava II and Dde I agarose genotypes, respectively (K-S test); shad from the Connecticut River could be differentiated from the combination of shad found in all other rivers and from the Chowan River/Albermarle Sound alone by significant differences of frequencies of shad mtDNA by significant differences of frequencies of shad mtDNA Dde I agarose genotypes (goodness-of-fit) and from shad from the St. Johns River and the Chowan River/Albermarle Sound by significant differences of frequencies of shad mtDNA Rsa I and Dde I agarose, and Hinf I and Dde I agarose genotypes, respectively (K-S test); shad from the Miramichi could be distinguished from shad from all other rivers except the Rappahannock by significant differences of frequencies of shad mtDNA Rsa I genotypes (K-S test); shad from the Delaware, Hudson, and Connecticut rivers had the greatest frequencies of the AAAAAAA composite mtDNA genotype; shad from the Hudson, Connecticut and Anapolis Rivers could be distinguished from those in other rivers by significant differences of frequencies of shad mtDNA EcoR V genotypes (X^2, Bentzen 1989); and shad from rivers in the St. Lawrence and Richelieu rivers could be differentiated from shad from all other rivers by significant differences of frequencies of shad mtDNA Hae II B genotype (X^2, Bentzen 1989).

The only significant direct correlation of latitude with genotype appeared to be with Dde I Genotype B, which decreased in frequency from south to north. This enzyme could have the potential to distinguish between stocks at different latitudes. Bentzen et al. (1989) also reported a cline in the frequency of mtDNA length genotypes (revealed with Kpn I and Sal I), with the largest mtDNA length

genotype found only at the southern end of the shad's range.

MtDNA molecular diversity

A. Intraspecific diversity

A large number of mtDNA fragments were detected that exhibited differing electrophoretic mobility, primarily on polyacrylamide gels. It is possible that these represent minor size variants due to differential copy numbers of short repetitive mtDNA sequences which may only be resolved in relatively small fragments. Support for this hypothesis comes from the earlier results of Bentzen et al. (1988) who detected multiple copies of short (less than 40 bp) repetitive mtDNA fragments in shad which probably mapped to the D-loop region. Both Densmore et al. (1985) and Wirgin et al. (1990b) observed small mtDNA fragments in lizards and striped bass, respectively, which differed in electrophoretic mobility on polyacrylamide gels. Although the molecular basis for this variation was not characterized.by Densmore et al. (1985), Wirgin et al. (1990b) attributed the variation to 130 bp tandem repeats. Concordance in molecular weight differences among these mtDNA samples was not seen when using a battery of different restriction endonucleases, an approach which can be used to unequivocally identify large mtDNA length variants (Avise et al. 1989). However, given the small size difference among these polymorphic mtDNA fragments, and the potential for these variations to go undetected in moderate to large mtDNA fragments, these results are not unexpected.

The hypothesis can not be disregarded that some of these polymorphic fragments harbor base substitutions that alter the conformation of these mtDNA fragments sufficiently to affect their electrophoretic mobility. Singh et al. (1987) also reported small human mtDNA fragments which differed in electrophoretic mobility on polyacrylamide gels. Cloning of these fragments and sequence

analysis revealed that these human mtDNA fragments contained non-restriction site base substitutions.

Heteroplasmy, which is the presence of two or more different size classes of mtDNA in the same individual, was detected with Dde I and Rsa I as revealed on polyacrylamide gels. The genotypes revealed were all length variants. Seven percent of the Dde I acrylamide-tested samples were heteroplasmic for two or three different combinations of genotypes. Of samples tested with Rsa I, 17% were heteroplasmic for different combinations of two genotypes. As depicted in Table 8, only one sample is heteroplasmic for both a Dde I acrylamide and Rsa I genotype. These data support the hypothesis that the heteroplasmy arose from independent events, a hypothesis also shared by Bentzen et al. (1988) (see below). Heteroplasmy has also been found in other animals, including the American shad (Bentzen et al. 1988). The mechanism of the occurrence of heteroplasmy is unknown, but it has been hypothesized that misalignment before replication could precede the inheritance of both length polymorphism and heteroplasmic molecules (Buroker et al. 1990).

Only one definitive restriction enzyme site polymorphism was detected (with EcoR I). This polymorphism took the form of the largest fragment in Genotype A (9973 bp) gaining a new restriction site as evidenced by two new fragments in Genotype B (7315 bp and 4500 bp) that add up to the same size as this large fragment in Genotype A. Of 257 shad mtDNA samples digested with this enzyme, all but 11 samples revealed Genotype A. Ten of the eleven Genotype B samples were located in the St. Johns River; one was found in the Hudson River.

I detected 98 different composite mtDNA genotypes of 161 shad tested (61% of the fish were different from each other). All diversity (with the exception of the EcoR I restriction site variant) resulted from the presence of

electrophoretic mobility variants. The St. Johns River had the highest restriction pattern diversity value, h = 0.975, as well as the largest number of genotypes unique to a single river.

Since more than half of the shad tested had unique composite mtDNA genotypes, an attempt was made to see if there was any concordance between the shared genotypes and geography with a phylogenetic analysis (see next section) Thus F and p values were calculated only for the 15 genotypes shared among rivers. P values ranged from 0.00057 to 0.0042 between each pair-wise comparison.

How did the results of this study compare to those of Bentzen and co-workers (1988 and 1989)? This approach differed in that it focused on the use of predominantly 4- and 5- base- cutting restriction endonucleases which produced many more fragments in shad than the use of predominantly six base-cutting restriction endonucleases. Bentzen et al. (1988 and 1989) sampled only 420 bp or 2.3 % of the shad mtDNA molecule whereas the enzymes in this study sampled 603 bp or 3.3% of the shad mtDNA molecule.

Restriction site variation was noted in 88% of the Bentzen et al. (1988 and 1989) samples, unlike the only 4% restriction site variation noted with EcoR I in my study. Bentzen (1988 and 1989) did not detect fragments of less than 500 bp because of the methodology employed, whereas most of the length polymorphisms I noted differed from each other by a range of from 1000 to 50 bp.

Heteroplasmy was observed for a Sal I restriction site polymorphism (4% of samples tested) and for a variant in the length (1 or more 1500 bp tandem repeats) of the mtDNA molecule (12% of samples tested) (Bentzen et al.1988). These percentages are similar to my observations (7% for Dde I acrylamide data and 17% for Rsa I data). Bentzen et al. (1988) also hypothesized

independent evolution of heteroplasmy as the longer shad mtDNA length variant they found also displayed the two different Sal I site polymorphisms. However, Bentzen et al. (1988) did not conjecture about the timing of the two events. Their data could indicate that length variation preceded base substitution or, alternately, that base substitution occurred twice.

Many fish in my study contained unique composite mtDNA genotypes (98 out of 161 fish tested or 60%). Bentzen et al. (1989) detected only 26 composite mtDNA genotypes of 243 fish tested or only 11% uniqueness. The most common genotype of Bentzen et al's. (1989) occurred in 90 or 38% of the shad tested whereas the most common genotype in my study was detected in only 7% of the fish. This indicates that I detected a greater amount of heterogeneity among and within shad populations.

Bentzen et al. (1989) detected overall levels of mtDNA sequence divergence among shad (mean p based on comparisons of all composite mtDNA genotypes found = 0.002) that are considered low compared to most p values detected for other intraspecific marine but about the same order of magnitude as p values detected for diadromous fish comparisons, as cited in Gonzalez-Villasenor and Powers (1990). My p values for pair-wise comparisons between fifteen shared shad composite mtDNA genotypes are similar in magnitude to those of Bentzen et al's. (1988) and range from 0.00057 to 0.0042.

As reported in the Gonzalez-Villasenor and Powers (1990) study, one of the the lowest mean p value between intraspecific comparisons of marine fish was 0.004 for Opsanus beta (Avise et al. 1987) and the highest was 0.029 for Clupea harengus (Kornfield and Bogdanowicz 1987) Mean p values of 0.002, 0.001 and 0.008 were reported for intraspecific mtDNA comparisons of the anadromous Onchorhynchus kisutch (Thomas et al. 1986), and the

catadromous Anguilla rostrata and Anguilla anguilla (Avise et al. 1986) respectively.

Low mtDNA divergence was also reported by Chapman (1987; 1989) and Wirgin et al. (1989; 1990b) (0.0004) between populations of striped bass from many of the same spawning rivers in this study. There was also a total absence of mtDNA diversity among anadromous Atlantic tomcod (Microgadus tomcod) from three spawning rivers within this same NW Atlantic area (Wirgin et al. 1990a). Somewhat higher levels of mtDNA sequence diversity were reported in populations of anadromous cisco (Coregonus artedii) from the James-Hudson bays (mean p = 0.0052) (Bernatchez and Dodson 1990). These workers hypothesized that post-glacial colonization of these systems by cisco from genetically divergent refugia may account for the current moderate levels of mtDNA diversity.

Comparisons of mtDNA genotypes in intraspecific fresh water and estuarine fishes yielded mean p values of up to a magnitude or greater than those of the marine, anadromous, and catadromous fish. This might be expected because, as these fish do not migrate, there would be more isolation of divergent mtDNA genotypes, which would contribute to higher p values.

Base sequence differences revealed by four base cutting enzymes were also successful at differentiating stocks of land-locked or anadromous Finnish Atlantic salmon and Swedish salmon from different drainage systems (Gyllensten and Wilson (1987) both in Davidson et al. 1989) whereas six-base cutting enzymes were not successful (Birt et al. 1986; Palva 1986) in differentiating allopatric populations of anadromous and non-anadromous salmon in Newfoundland (Birt et al. 1986) or Finland (Palva 1986).

fishes differ by as much as an order of magnitude., which could be indicative of when the species split in geologic time. An example of one of the lowest reported divergence values is 0.027 for Salmo clarki and S. gairdneri (Wilson et al. 1985) and one of the highest is 1.000 for Lepomis cyanellus and L. marginatus (Avise and Saunders 1984).

Shared composite mtDNA genotypes

Figure 11 shows how 15 composite shad mtDNA genotypes were shared among or between rivers. For purpose of this discussion, the rivers as displayed in the diagram were "adjacent" to each other, even though it is understood that there were frequently geographically intervening rivers not sampled by this survey. "Continuous" or "discontinuous distributions" were used in the same spirit. Thirteen of the 15 genotypes had discontinuous distributions. Four genotypes were shared by two rivers that were not geographically proximal. There were no genotypes shared between the St. Johns and the Miramichi (the two most distant rivers) nor between the Rappahannock and the Delaware (adjacent rivers) in this survey. The latter may represent a historic break in the distribution of shad. South of the Rappahannock, no genotypes was shared by more than two geographically adjacent rivers, whereas north of the Rappahannock River, four genotypes were shared by at least three continuously distributed rivers. This could indicate that there is more mixing of shad stocks among northern rivers.

Phylogenies relating shared genotypes

Two approaches were undertaken in an attempt to see if there was any concordance between molecular relationships as detected by shared genotypes and geographic proximity of these populations. An assumption of

the distance approach was that site data was used, when, in reality, fragment data was used. Another assumption was that all fragments were homologous. Parsimony analysis used only fragments that were assumed to be synapomorphic (and homologous) in each genotype. (These fragments may have actually been sympleisiomorphies.)

Homology was generally used to mean two different concepts that may or may not be mutually exclusive. A "homologous fragment", for example, has the same terminal sequence (i.e. restriction site) as the fragment to which it is "homologous". The assumption was also made that homologous fragments were located in the same places in each molecule. Homology in a broader context refers to common origin of a character in an immediate common ancestor that is shared by two descendents (Wagner 1989).

The unrooted distance phenogram separated the 15 shared genotypes into 3 clusters, the middle one of which contained genotypes that were shared by the St. Johns and the other rivers. No genotypes present in the St. Johns River were found in the other two clusters. When this phenogram was rooted with the hickory shad p values, the clusters broke out in a slightly different way. The hickory shad was rooted approximately midway between a cluster on the bottom of the figure , which contained the three most widespread genotypes across rivers (some of which were contiguous) and a top cluster, the genotypes of which had a more restricted distribution. The hypothesis could be evoked that the hickory shad mtDNA genome is equally divergent from widespread and restricted mtDNA genotypes. If this were the case then the genotypes in one cluster are probably not "older" than those in the other, but instead some other physical or biological mechanism may have influenced the distribution of the genotypes.

The three most parsimonious unrooted HENNIG86 cladograms indicated

that there was no concordance with geography with the unrooted cladograms. However, the HENNIG86 cladogram rooted with hickory shad was interpreted as showing that the hickory shad was more closely related to those fish which did not display genotypes in common with the St. Johns River. There was no other geographic partitioning with the other genotypes.

Genetic differentiation and gene flow

Gene flow is often equated with Nm, where N = the population size and m = the migration rate. A gene flow greater than 1 is considered to be moderate and Nm less than 1 is low (Slatkin and Maddison 1989). Slatkin (1985) did computer simulations that relate gene flow to the number of distinct (rare) alleles within a population. (In later papers, Slatkin (1989), Slatkin and Maddison 1989) reported that alleles can be equated with mtDNA genomes). These simulations depict an approximately inverse linear relationship between gene flow and the average frequency of rare alleles. Slatkin (1985) also uses the term "private" alleles which is synonymous with my term "unique" alleles to describe alleles found only in one deme or subpopulation. In this case, a deme would be an individual river. Thus, the higher the average frequency of rare alleles in a river, the more genetically distinct this subpopulation is from the others, and thus the lower the amount of gene flow between rivers.

However, Slatkin (1985) mentioned that gene flow results obtained in this manner are subject to bias based on population sample numbers and that, upon further sampling, the number of rare alleles will either increase or decrease proportionately, which would affect the amount of gene flow determined. Also, genotypes considered to be private could be detected in previously unsampled intervening rivers.

An island model of dispersal or gene flow is assumed for Slatkin's

simulations (equal probability of migration from each population to every other population) as opposed to a stepping-stone model (greater likelihood of migration between adjacent populations) (Slatkin 1985; Slatkin 1989; Slatkin and Maddison 1989). Other assumptions include a constant population size that is equal among demes and neutral selection of alleles. Slatkin's (1989) simulations (at the present time) can only handle a limited number of migration events and are costly as they require a long time to run.

If the Slatkin (1985) analysis could be applied to the mtDNA data in this study the large number of rare composite mtDNA genotypes within rivers was indicative of genetic differentiation. The river with the greatest number of distinct composites (the St. Johns--90%) had the lowest amount of gene flow (according to Slatkin's simulations (1985)). A range of 35-60% of the composites in the other rivers were unique, so thus a fair degree of genetic isolation must be assumed among the rivers.

However, shared composite genotypes among rivers represented gene flow among rivers. There was an absence of substantial gene flow among shad spawning systems which were the most geographically distant (the St. Johns and the Miramichi) as evidenced by a lack of common composites.

Why are some populations more genetically diverse than others? Possible answers are: (1.) The more diverse populations could be older and thus have had more time for divergence of mtDNA genotypes than younger populations. (2.) There may be more geographic isolation (such as many tributaries) within certain individual river systems leading to vicariance or restricted gene flow within those rivers. This could allow for more divergence of genotypes over time or even microhabitat selection of mtDNA genotypes in these rivers, as opposed to those in rivers with a more uniform hydrography. (3.) There may be larger population sizes in some rivers which could reflect or

lead to the greater diversity of mtDNA genotypes in these rivers when compared to mtDNA genotypes in rivers with smaller population sizes. (4.) All populations could have originally had the same magnitude and types of diversity, but population bottlenecks or extinctions may have lead to a smaller gene pool of mtDNA genotypes in some populations. (5.) Founder effects could produce new populations with a disproportionate distribution of mtDNA genotypes from a previous gene pool. These new populations could be more or less genetically diverse than the old populations. (6.) Selection may actually favor some mtDNA genotypes over others and lead to greater homogeneity in some populations then others. Some Drosophila populations exhibit non-neutral mtDNA genotypes (MacRae and Anderson 1988). (Selection may actually mask gene flow (Slatkin 1985), as immigrant mtDNA genotypes may actually be unadapted to their new habitat and may then go extinct.)

Possible explanations of diversity within populations

Assuming that ancestral shad populations were restricted to the southeastern US during glacial maxima (Fodor 1981), these older southern populations would be expected to exhibit more variability because they have had additional time to undergo mtDNA alterations. The only consistency in this idea is displayed within shad mtDNA from the St. Johns River, which had the highest restriction pattern diversity value, $h = 0.975$, as well as the largest number of composite mtDNA genotypes unique to a single river. There was no apparent south to north decline in restriction pattern diversity values.

This is in contrast to Bentzen and co-workers (1989) who found the heterogeneity of mtDNA genotypic frequencies to be greatest within rivers north of the Delaware. They attributed this finding to founder effects resulting from post-glacial colonization of these rivers from southern refugia.

Shad populations violate the assumption of a constant population size that is equal among demes, as population sizes vary among and within rivers. Population size as measured in pounds X 10^3 of spawners indicate from four to ten-fold fluctuations for several populations, including the St. Johns, Altamaha, Chowan, Delaware, Hudson, and Connecticut rivers, as recorded since 1950 to the late 1970's and 1980's (Gibson 1988). For example, the largest number recorded for the Hudson river was 3043 pounds X 10^3 of spawners in 1980. Because of the large population size fluctuations and the absence of data for some of the rivers and particular years in which I sampled shad, it was difficult to relate numbers of genotypes found to population size. If the St. Johns River population is indeed smaller than the rest of the populations, the afore-mentioned hypothesis that greater genetic differentiation would be found in older (southern) populations is enhanced. The size of the river is thought to be responsible for some of the among-river variation in population abundances (Gibson et al.1988) and recruitment variation could explain some of the fluctuations that could have lead to reductions or local extinction events in within-river populations of shad.

Density-dependent and density-independent factors affect shad recruitment and adult population size (and thus population fluctuations). Density-dependent factors (adult population size, cannibalism, competition, availability of prey and predation) and density-independent factors (river flow, water temperature, rainfall) are responsible in varying degrees depending on pre-or post-juvenile stage for levels of recruitment in the Connecticut River (Crecco and Savoy 1984, 1987; Savoy and Crecco 1988).

Annual year-class strengths from 1967-1984 were significantly inversely correlated with mean river flows and directly significantly correlated to river temperatures in the month of June, (the month during which most shad larvae

hatch in the Connecticut river), which supports the hypothesis that hydrographic factors (density-independent) influence recruitment, at least during the early larval stages.

Since there is poor correlation between these same hydrological factors and the abundance of eggs, older larval and juvenile stages, density-dependent factors probably have a greater effect on egg and later stages. There is also a significant positive correlation between mean feeding incidence (determined by the presence of food in the gut) on zooplankton in May and June (1979-1984) and year-class strength. However, there are inconsistent correlations between feeding incidence and zooplankton densities in the river, which the authors could not explain (Crecco and Savoy 1984). This could possibly be due to selective feeding by the shad or threshold levels of zooplankton.

Thus, there is a coupling of density-independent and density-dependent factors that probably influence shad abundance, at least in the Connecticut River. Stronger year classes appear to be more dependent on lower river flows, higher water temperatures and higher feeding incidence than are weaker year classes. Crecco and Savoy's work (1988) supports a hypothesis that a mixture of density-independent and density-dependent factors are predominant at influencing year-class strength in the early life of the shad, and density-dependent factors are more important in subsequent stages . Abundance of parental stock, however, was not correlated with adult recruitment of females (Crecco and Savoy 1984). Gibson (1988) purports, however, that there is probably a threshold adult stock level below which shad populations could be decimated, which supports an ultimate density-dependent mechanism (adult spawner abundance) that controls shad population levels.

Water quality conditions such as sewage loading and dissolved oxygen in

addition to density-independent factors are related to adult shad population abundance in the Potomac, Delaware and Hudson rivers (Summers and Rose 1987). These authors found that, in addition to river flow, shad abundance is negatively correlated with increased sewage loading in the Potomac, and with decreased dissolved oxygen levels in the Delaware and Hudson rivers. Unfortunately, the projected abundance models for this study are based on commercial fisheries catch statistics that have been recorded since the 1880's. Population statistics before that time were not recorded so there is only a 100 year history of population fluctuations. Abundance data are often extrapolated from catch/unit effort data, the effort of which is considered to be a constant. However, this constant is often violated, as fishing landings data are often subject to market fluctuations, such as a rise and fall in the demand for shad roe, and the cessation of shad fishing during World War II (Richkus and DiNardo 1984).

Life-history strategies could also account for population fluctuations. Greater fecundity of the southern shad appears to be balanced by the iteroparity of the northern fish, and thus fecundity across rivers is relatively constant in proportion to population abundance (Leggett and Carscadden 1978). Perhaps the greater fecundity of the southern fish may not make up for any detrimental effects to progeny of the single spawning and consequently less fish would be available to recruit into the next generation. The repeat spawners are available for subsequent spawnings and could possibly overcome the adverse effects of one spawning and this may indeed be part of the reason for the larger population sizes in some of the northern rivers.

The number of unique genotypes present in a population could be indicative of recent population sizes. A lower number of genotypes could indicate a population reduction based on the mechanisms outlined above.

Thus, mtDNA sampling could be used to estimate recruitment success in the absence of more direct means. Random sampling and neutral selection would have to be assumed.

Gene flow may be responsible for the inability to distinguish some populations of shad via certain frequency distributions of mtDNA genotypes. Infidelity between systems may be currently, or in the recent evolutionary past, substantial, or gene flow between these two systems could currently be minimal, but insufficient time has elapsed since their division for the buildup of genotypic frequency differences. Given that the impact of the most recent glacial event was only about 18,000 years ago (Fodor 1981), only slight differences in mtDNA genotypic frequencies might be anticipated between any two northern systems. For example, if one were to apply a rate of mtDNA change of 2% per million years (Brown et al. 1979), then a diversity value of approximately $p = 0.0004$ would be projected when comparing Hudson and Delaware shad, for example. The maximum mtDNA sequence diversity between striped bass from the same geographic area (Hudson River and Chesapeake Bay) is reported as $p = 0.0004$, a value which is consistent with these estimates (Wirgin et al. 1989b). It is possible that the use of additional restriction endonucleases on shad mtDNA samples may reveal additional informative polymorphisms and thus it would be premature to consider any two spawning systems as a single stock.

Distribution of shared genotypes

How can the distribution of the shared composite mtDNA genotypes be explained? Composite mtDNA genotypes shared by adjacent northern and southern rivers could represent straying. The dispersal mechanism is unknown with shad populations, but since spawning appears to be temperature

dependent (Leggett and Whitney 1972) it is probably more likely that dispersal would be by straying of a fish from one river to an adjacent river because of similar thermal regimes. However, since these fish are capable of migrating long distances, island dispersal should not be ruled out. If an island model was assumed with this data set, but dispersal really occurred through a stepping-stone model, gene flow would be underestimated because it would take more "steps" (migration events) to get to non-adjacent rivers.

Some infidelity could be related to the life-history strategies of the fish. Shad below 32 N are semelparous (spawn only once) whereas the proportion of shad that are iteroparous (spawn repeatedly) increases linearly with latitude north of Cape Hatteras, NC. (Leggett and Carscadden 1978). If age correlated with infidelity (not found in the literature) a larger proportion of repeat spawners could be correlated with a larger number of infidels which could explain the larger proportion of shared genotypes among the northern rivers.

Some of the disjunctions in mtDNA composite genotypes could be due to inadequate sampling, recent population restrictions in rivers in which those genotypes that were originally present were lost, drift and/or mutation with or without vicariance in rivers lacking the genotypes, straying of shad that "skipped" the river lacking the genotypes, and/or stocking by artificially propagated shad.

It is unknown as to what effect shad stocking programs in the 1870's and 1880's (U.S. Fish Commission Reports) had on disrupting homing ability of shad. Frequently shad were transported from one river and planted in another. Even though shad were placed initially in the Sacramento River in 1871, they spread to many rivers and now range from Baja, California to Kodiak, Alaska (see references in Leggett 1976). This might explain some of the lack of concordance between the shared composite mtDNA genotypes' pattern of

distribution and geography.

Comparison of this to other mtDNA studies on clupeids

Avise et al. (1989) also reported numerous composite mtDNA genotypes resulting from both restriction site polymorphisms and presumptive length variants when screening two collections of the marine clupeid fish menhaden (Brevoortia tyrannus/B. patronus). Thirty-two of 33 fish exhibited individual-specific composite mtDNA genotypes when tested with 18 five-and-six base cutting restriction endonucleases. These workers invoked the enormous effective population of menhaden (greater than 10^8) as a likely explanation for this extensive polymorphism. Kornfield and Bogdanowicz (1987) also reported substantial levels of mtDNA base sequence diversity and an absence of mtDNA length variants in another species of marine clupeid fish, the Atlantic herring (Clupea harengus). Current estimates of population size of Atlantic herring are also very large which may in part account for their substantial levels of base sequence diversity. The menhaden and herring data suggest a rapid rate of mtDNA base sequence change in the clupeid fish, which is probably comparable to that observed in primates and other animals (Brown et al. 1979). The finding of a lack of base sequence divergence in American shad, in combination with extensive electrophoretic mobility polymorphisms, suggests a rapid rate and recent generation of the latter form of variants.

Applications of mtDNA analysis to problems in shad biology

A. Stock identification

Since over 100 restriction enzymes are now currently commercially available, more enzymes could be tested on shad mtDNA in search of polymorphisms that are more river or region specific. Perhaps one enzyme is

better at detecting sequences of mtDNA that are region or river-specific in terms of energy usage related to mitochondria. These sequences could be isolated and amplified via cloning or PCR.

There are faster ways of extracting mtDNA, such as the use of vertical and fixed- angle ultracentrifuge rotors that require less spinning time for separation of mtDNA from nuclear DNA and rapid-prep methods of mtDNA extraction that do not require centrifugation (Maniatis et al.1989) and improved gel techniques such as southern blotting and electroelution that assure that homologous fragments are being compared.

American shad were introduced into the Sacramento-San Joaquin river system on the Pacific coast in 1871. These fish were transferred from New York, (Baird 1871) and through migration and straying now range from Baja, California to Kodiak Island, Alaska (see references in Leggett 1976). Could mtDNA analysis be used to differentiate shad from these rivers? Although it is doubtful because of the short 100 year time frame since these fish were introduced, founder effects may be assessed. If significant differences in frequencies of shad mtDNA genotypes genotypes between Pacific and Atlantic rivers were found, this could be an indication of founder effects, or that the divergence clock for shad would have to be modified. Homing is perhaps not as stringent as it might be on the Atlantic coast because these are introduced fish. Of twelve tags returned from 6000 tagged fish, nine were definitely from the river of origin, either the Feather River or the Upper Sacramento River (Stevens et al. 1987). If river-specific mtDNA genotypes were found, could this information be used to underscore quantity of homing and/or dispersal patterns in these fish?

Also it has been noted that there appear to be two separate populations in San Francisco Bay, one that migrates, and one that stays in the bay (Maxwell

Eldridge, personal communication). Can mtDNA analysis delineate these two sets of fish?

Recently, researchers have used a plethora of techniques such as nuclear DNA probes (Wirgin and Maceda 1991), nucleolar organizer regions (NOR) on chromosones (Legrande and Grady 1991), immunoassays, logistic regression, and fatty acid composition (Grahl-Nielsen and Ulvund 1990; Waldman 1991) in an effort to differentiate populations of various fishes. Three nuclear probes (of 16 tested) differentiated striped bass populations from the Atlantic versus the Gulf of Mexico and South Carolina versus the Hudson-Chesapeake (Wirgin and Maceda 1991). In each of three populations of the madtom (Noturus elegans) a unique NOR (silver and/or CMA) condition was observed (Legrande and Grady 1991). The fatty acid composition technique might particularly be an appropriate technique to use with shad because the amount of fat depleted during migration correlated directly with distance traveled, and thus river of origin (Glebe and Leggett 1981). Perhaps a variety of techniques tested in conjunction with mtDNA analysis would be the best way to tackle stock discrimination.

B. Hatchery restoration

Hynes et al. (1989) used mtDNA as a genetic marker to differentiate between natural and hatchery-reared stocks of brown trout (Salmo trutta). Since morphometric measurements could not distinguish wild from cultured juvenile American shad (Johnson and Loesch 1986), mtDNA analysis might be used to make this distinction. A definitive marker could then be used to assess the percentage of survival/contribution of hatchery fish to the populations.

Because of serious depletions of shad, the Susquenhanna River in Pennsylvania and the Merrmiac River in New Hampshire are currently being

stocked with shad from the Hudson and Connecticut rivers respectively (Boyd Kynard, personal communication). The genetic integrity of the different populations of shad using mtDNA analysis could be assessed in an attempt to determine whether stocking programs are prudent.

C. Interspecific diversity

Originally shad on the Atlantic and Gulf of Mexico coasts were differentiated into three species based on gill raker counts (A. sapidissima, A. alabamae and A. ohiensis) (Leggett 1976) but this has been disputed so that now only the former two are recognized as being separate species (Hildebrand 1963). MtDNA analysis could be used to differentiate these closely related species and perhaps lend itself to dividing subspecies. A project is currently underway to study whether Alabama shad (Alosa alabamae) can be differentiated from American shad using mtDNA analysis (Wirgin, personal communication).

Conclusions

The shad, as do other anadromous fish, present an unusual management problem in that it is difficult to ascertain what the "unit stock" for these fish is. Shad do home to natal rivers (tagging studies) and there is river-specific population genetic differentiation in the form of unique single-enzyme and composite mtDNA genotypes. There were significant differences in frequency distributions in some American shad mtDNA genotypes between all rivers tested. This indicates that the shad populations in this study could be differentiated. Thus the populations in these rivers should be managed as separate stocks.

Two mtDNA markers, the EcoR I B genotype and the Dde I polyacrylamide

F genotypes could be used to assign with a high degree of accuracy the river of origin of a St. Johns or a Miramichi fish in a mixed marine sample (pending program development by Crittenden). Intervening rivers did not display such river-specific markers, so stock composition analysis of the ocean shad fishery is limited when using the enzymes in this study for mtDNA analysis.

Even though the populations in these rivers should be managed as separate stocks, representatives of all populations of shad are found in the Bay of Fundy, the St. Lawrence estuary, and off the Newfoundland and Labrador coasts during the summer months (Dadswell 1987). Also, there was evidence of gene flow between populations in the form of shared mtDNA composite genotypes. Thus, all populations of shad should be managed as a unit, especially during ocean migrations.

Summary

MtDNA diversity was evident in the form of length variants and a single restriction site polymorphism. Four- and five- base cutting restriction endonucleases detected more heterogeneity in shad populations than did 6-base cutting enzymes (many fish contained a rare composite genotype); however, this did not translate into higher genetic divergence values than Bentzen et al's. (1989) study. Divergence values were compared to those of other marine, estuarine, anadromous and catadromous and fresh-water fish. Interspecific divergence values were an order of magnitude lower than what would be expected between A. sapidissima and A. mediocris. Fish containing shared composite mtDNA genotypes were related phylogenetically with mtDNA analysis, with the distance phenogram producing some concordance between geographic location and genotype. Genetic differentiation and gene flow were discussed and explanations of dispersal patterns of shared composite

genotypes were explored. Comparisons were made of this study to others with clupeids. Finally, other potential applications of mtDNA analysis for solving problems in shad biology were examined.

Table 1. Shad collection data.

River	Date	Sample number	Total number	Sex	Std. or fork length in mm	Total length in mm	Weight in g	Gear
St. Johns, FLA	3/4/85	297-308	12	F				electroshock
	3/6/86	394-407	14	4F:10M	297-438	372-536	468-1550	hook & line
(Mullet Lake)	2/17/87	527-545	19	F	420-475		650-1250	
			45					
Altamaha, GA	3/13/86	408-422	12	F	430-560	480-600		
	2/23/87	546-560	15	F	460-550		1450-2300	gill net
			27					
Santee-Cooper, SC	3/29/86	423-434	11	7M:4F	430-460	460-480		electroshock
(1 hickory shad)	&4/3/86	428	1	F	430	480		
diversion canal (all)	3/28/87	561-574	14	F	545-595		1920-3060	electroshock
			26					
Chowan/Albermarle	4/10/86	436-454	19	14M:5F	350-550	380-600		gill net
Sound, NC	4/1/87	578-592	15	F	410-525		1700-2700	
			34					
Rappahannock, VA	4/21/87	598-617	**20**	F				
Delaware, NY/NJ/PA	5/26/85	345-346	2	1F:1M	560-570	580-610		hook & line
Ten-Mile River (no361)	6/20/85	359-370	11	6F:5M	445-550	500-610		drift net
Port Jervis, NY	5/4/86	470-476	7	4M:3F	480-530	530-580		hook & line
Lambertville, PA	5/1/87	620-639	20	F	460-530		1590-2600	haul seine
			40					
Hudson, NY/NJ	5/23/84	198	1	F				
GW Bridge	5/9/85	324-329	6	F	510-540			gill net
Nyack	5/16/85	334-339	6	F	460-500			gill net
Claverack	5/18/85	340-344	5	F	515-550			haul seine
Nyack	5/1/86	458-469	12	F	480-570		2300-2700	gill net
Nyack	5/3/87	640-659	20	F	470-560	520-560		gill net
			50					

Table 1. Shad collection data.

River	Date	Sample number	Total number	Sex	Std. or fork length in mm	Total length in mm	Weight in g	Gear
Connecticut, MA	6/21/84	216-219	4	3F:1M				elevator
Holyoke Dam (all)	6/11/85	347-358	12	F	460-560	505-610		elevator
	6/7/86	477-500	24	F	480-550	530-620	770-1140	elevator
	5/15/87	661-678	18	F	500-550	560-600	1000-1400	elevator
			58					
Miramichi, NB Canada Millbank	6/3/87	681-697	17	F	451-525		1200-2000	

Table 2. MtDNA fragment sizes in base pairs for each shed single-enzyme mtDNA genotype.

Ava II genotypes:

A	B	C	D	E	F	G	H	I	J	K
2371	2395	2375	2390	2390	2390	2404	2536	2366	2385	2385
2134	2134	2158	2180	2180	2150	2163		2148	2171	2158
			1900				1940			
1870	1867	1881	1800		1870	1881		1839	1867	1881
1718	1718	1733	1630	1690	1737	1730	1793	1688	1737	1737
1677	1710	1730		1690	1648	1730	1560	1638	1737	1737
		1578			1515	1507		1556	1532	
								1486		
1462		1445								1477
	1425		1430	1400	1433		1405	1407	1433	
1190	1194		1180	1200	1216	1218	1279	1194	1209	1216
				1100						
958	952		955	920	962	962	954		956	964
864	867	872	861	875	882	878	877	846	876	885
				715						
			583							
		557								
529	529	529	522	508	583	583	581	590	564	589
423	418	430	421	418	436	435	458	415	423	446
414	414	420	414	414	414	414				
399	399	400	403	399	399	399				
378	378	380	386	378	378	378				
351	351		354	351	351	351				
325	325	330	334	325	325	325				
		325								
300	300	302	312	300	300	300				
287	287	280		287	287	287				
260	260	250	259	260	260	260				
248	248	240	226	248	248	248				
Total bp 18158	18171	18215	18540	18048	19784	18453				

Data for Ava II genotypes A-G was obtained from both agarose and polyacrylamide gels. Polyacrylamide data for Ava II genotypes H-J is not available, thus fragment numbers will not add up to the unit size of the molecule. Totals are given for EcoR I data. Dde I agarose and Hinf I polyacrylamide data do not add up to the unit size. Further resolution is needed to visualize all fragments but was not warranted in this study.

Spaces indicate a missing fragment for that genotype that is present in another genotype.

Table 2. MtDNA fragment sizes in base pairs for each shed single-enzyme mtDNA genotype.

Ava II genotypes:			Dde I genotypes: agarose data					
L	M	N	A	B	C	D	E	F
	2857							
2457		2437	1600	1600	1600	1600	1600	1600
2211	2223	2184			1430		1500	
								1260
1872	1869	1862		1015		1015	1015	1015
1710	1717	1701	810	810	810	810	810	810
1642	1655	1641	810	810	810	810	810	810
		1564	735	735	735		735	735
	1497						695	
			650	650		650		650
1455		1438	595	595		595	595	595
1248	1243		540					
			520	520	520	520	520	520
890	890	934						
793	793	851						
		564						
		434						

Table 2. MtDNA fragment sizes in base pairs for each shad single-enzyme mtDNA genotype.

Dde I genotypes: polyacrylamide data

A	B	C	D	E	F	G	H	I	J	K	BC
513	512	494	509	513	506	506	506	508	511	506	510
510	512	480	503	510	488	487	487	501	508	487	506
491	496	458	479	491	469	464	464	480	481	464	475
480	496	458	479	480	459	460	460	470	481	460	470
455	449	449	446	455	434	434	434	446	451	434	445
455	449	449	446	455	434	434	434	446	451	434	445
422	423	427	410	422	414	410	410	420	421	410	440
						404					
394	394	395	395	394	388	385	385	398			395
	377							362			375
352	359	387	359	352	346	345	345	359	349	345	350
				335							
327	331	326	322	327	325	326		325	326	326	330
315	319	317	312	315	316	315	315		316	315	320
296	295	290	289	296	292	292	292	299	293	292	295
296	295	282	288	296	292	287	287	299	290	287	295
								289			
284	288	279	280	284	284	281	281	288	280	281	280
284	288	279	280	284	284	281	281	288	280	281	280
271	275		253	271	261	267	267	268	260	267	
271	275		253	271	261	267	267	268	260	267	
245	251	238	238	245		242	242	242	238	242	235
227	236	220	223	227	220	223	223	223	225	223	222
202	208	199	204	202	200	200	200	199	202	200	200
202	208	199	204	202	200	200	200	199	202	200	200
187	190	184	184	187	185	185	185	186	182	185	177
175	183	174	173	175	175	173	173	174	170	173	170
175	183	174	173	175	175	173	173	174	170	173	170
169	175	166	168	169	169		161	166	165		165
160	169	158	161	160	163	159	159	157	158	159	157
156		153	157	156	156	153	153	158	155	153	155
150	153	149	151	150	150	149	149	148	145	149	145
136	143	132	135	136	136	134	134	131	125	134	135
133	139	130		133	134	132	132		125	132	130
	132				110					115	
96	100	94	92	96	91	95	95	92	85	95	92
91		91		91	87	88	88	90		88	92
75	72		71	75	74	71	71	72	75	71	71
69			61	69	71	60	60		60	60	63
					67						

Table 2. MtDNA fragment sizes in base pairs for each shad single-enzyme mtDNA genotype.

Dde I genotypes: polyacrylamide data						EcoR I genotypes:	
BE	BG	BGF	DE	EF	EH	A	B
518	508	506	507	508	508	9973	
506	497	488	498	496	493	7315	7315
484	478	469	476	474	472		5097
484	470	459	464	463	472		4220
474	447	434	443	438	442	1377	1377
474	447	434	443	438	442		
	429						
444	422	414	416	415	420		
		402					
398	391	388	394	393	391		
376	356	364					
351	347	346	348	349	348		
335			334	328	335		
326	328	325	327	325			
318	317	316	316	315	317		
305	292	292	288	294	293		
305	288	287	290	292	292		
294	280	282	280	284	284		
286	280	282	280	284	284		
269	253	261	253	261	268		
269	253	261	253	261	268		
246	238		238		243		
228	223	226	225	220	220		
204	204	200	202	200	200		
204	204	200	202	200	200		
189	184	189	182	185	185		
178	173	177	170	175	175		
178	173	177	170	175	175		
170	168	170	165	169	169		
161	161	162	158	163	163	18665	18009
149	157	154		156	156		
144	151	144	145	150	150		
134	135	123	125	136	136		
132	131	123	125	134	134		
		91	85	93	93		
		87					
		74	75	77	77		
		71	60	67	67		

Table 2. MtDNA fragment sizes in base pairs for each shad single-enzyme mtDNA genotype.

Hinf I genotypes: agerose data				Hinf I genotypes: polyacrylamide data	
A	B	C	D	A	B
		1327	1327		
1250	1270			1353	1353
	1015	1041	1031	1050	1050
			1011	1030	1030
949				790	790
932	932			740	740
880	880	828		710	710
		781		680	680
		781		640	640
705	705	705		530	530
				515	515
				500	500
				480	480
				470	470
				430	430
				428	428
				385	385
				380	380
				365	365
				355	355
				310	310
				295	295
				265	265
				230	230
					225
				215	215
				188	188
				175	
				165	165
				162	162
				157	157
				128	128
				112	112
				98	98
				88	88
				82	82

Table 2. MtDNA fragment sizes in base pairs for each shad single-enzyme mtDNA genotype.

Rsa I genotypes:

A	B	AB	C	AC	BC	D	BD	E	F	G	H
525	525	525	525	525	525	525	525	525	525		
495	495	495	495	495	495	495	495	495	495		
460	460	460	460	460	460	460	460	460	460		
460	460	460	460	460	460	460	460	460	460		
410	410	410	410	410	410	410	410	410	410		
410	410	410	410	410	410	410	410	410	410		
400	400	400	400	400	400	400	400	400	400		
395	395	395	395	395	395	395	395	395	395		
395	395	395	395	395	395	395	395	395	395	395	399
	365	365			365		365				
355		355	355	355	355			355	355	343	
										338	342
			330	330	330						
325	328	325		325						322	323
								315		314	
305	305	305	305	305	305	305	305	305	305	305	309
285	285	285	285	285	285	285	285	285	285		
280	280	280	280	280	280	280	280	280	280		
275	275	275	275	275	275	275	275	275	275		
260	260	260	260	260	260	260	260	260	260		
240	240	240	240	240	240	240	240	240	240		
240	240	240	240	240	240	240	240	240	240		
225	225	225	225	225	225	225	225	225	225		
205	205	205	205	205	205	205	205	205	205		
200	200	200	200	200	200	200	200	200	200		
182	182	182	182	182	182	182	182	182	182		
168	168	168	168	168	168	168	168	168	168		
									145		
									105		
122	122	122	122	122	122	122	122				
90	90	90	90	90	90	90	90	90			

Table 3 a. Frequencies of individual restriction enzyme mtDNA genotypes for each river.

EcoR I Genotypes				Hinf I Ag Genotypes					Hinf I Ac Genotypes		
	A	B		A	B	C	D	Totals	A	B	Totals
River			Totals								
STJ	27	10	37	25	8	0	0	33	36	1	37
ALT	25	0	25	7	2	0	0	9	11	0	11
SAN	22	0	22	16	2	1	1	20	14	3	17
CHO	27	0	27	119	3	0	0	22	11	10	21
RAP	16	0	16	10	3	0	0	13	15	0	15
DEL	28	0	28	10	12	0	0	22	19	8	27
HUD	38	1	39	18	8	0	0	26	29	6	35
CON	48	0	48	28	9	0	0	37	44	1	45
MIR	15	0	15	7	5	0	0	12	15	0	15
Totals	2246	11	257	140	52	1	1	194	194	29	223

Rsa I Genotypes													
	A	B	C	D	AB	AC	BC	BD	E	F	G	H	Totals
River													
STJ	4	6	17	1	1	0	4	1	0	0	0	0	34
ALT	8	1	3	0	0	0	0	0	0	0	0	0	12
SAN	4	11	2	0	0	0	0	0	0	0	1	0	18
CHO	5	13	2	0	0	0	0	0	0	0	0	0	20
RAP	10	2	1	0	0	0	0	0	1	0	0	0	14
DEL	14	6	1	0	1	2	0	0	0	1	0	0	25
HUD	21	7	4	0	1	1	0	0	0	0	0	0	34
CON	26	8	8	0	3	0	0	0	0	0	0	1	46
MIR	11	1	0	0	0	0	0	0	0	0	0	0	15
Totals	103	55	38	1	6	3	4	1	1	1	1	1	218

Table 3 a. Frequencies of individual restriction enzyme mtDNA genotypes for each river.

Ava II River	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Totals
STJ	11	15	3	0	0	1	2	0	1	1	0	0	0	0	34
ALT	6	4	0	0	0	2	0	0	0	0	1	3	0	0	16
SAN	13	4	0	0	0	0	0	0	0	0	0	2	0	0	19
CHO	9	3	0	0	0	0	0	0	0	0	0	7	1	1	21
RAP	8	3	0	0	0	0	0	0	0	0	3	0	0	0	14
DEL	11	12	0	0	0	0	1	1	0	0	0	0	0	0	25
HUD	21	12	0	0	1	0	1	0	0	0	0	0	0	0	35
CON	23	15	0	0	0	0	0	0	0	0	0	6	1	0	45
MIR	7	7	0	1	0	0	0	0	0	0	0	0	0	0	15
Totals	109	75	3	1	1	3	4	1	1	1	4	18	2	1	224

Dde I Ag River	A	B	C	D	E	F	Totals
STJ	12	22	0	1	0	0	35
ALT	9	7	0	0	0	0	16
SAN	15	3	0	0	1	0	19
CHO	6	13	1	0	0	1	21
RAP	8	7	0	0	0	0	15
DEL	17	7	2	0	0	0	26
HUD	22	11	0	0	0	0	33
CON	38	3	1	0	0	0	42
MIR	11	0	4	0	0	0	15
Totals	138	73	8	1	1	1	222

Dde I Genotypes

Ac River	A	B	C	D	E	F	G	H	I	J	K	BC	BE	BG	BGF	DE	EF	EH	Totals
STJ	25	1	0	1	4	0	0	0	0	1	0	0	0	0	0	2	0	0	34
ALT	10	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13
SAN	10	0	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14
CHO	14	0	2	1	2	0	0	1	0	0	0	0	0	0	0	0	0	0	20
RAP	10	0	2	1	1	0	0	0	0	0	0	0	0	1	0	1	0	0	16
DEL	18	0	3	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	24
HUD	26	3	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	31
CON	34	0	1	1	2	0	1	0	1	0	1	0	1	1	0	0	0	1	43
MIR	4	0	1	0	0	6	0	0	0	0	0	0	0	0	1	0	1	0	13
Totals	151	4	14	7	10	6	1	1	1	1	1	2	1	3	1	3	1	1	209

Table 3 a. Frequencies of individual restriction enzyme mtDNA genotypes for each river.

Rsa I Genotypes

River	A	B	C	D	AB	AC	BC	BD	E	F	G	H	Totals
STJ	4	6	17	1	1	0	4	1	0	0	0	0	34
ALT	8	1	3	0	0	0	0	0	0	0	0	0	12
SAN	4	11	2	0	0	0	0	0	0	0	1	0	18
CHO	5	13	2	0	0	0	0	0	0	0	0	0	20
RAP	10	2	1	0	0	0	0	0	1	0	0	0	14
DEL	14	6	1	0	1	2	0	0	0	1	0	0	25
HUD	21	7	4	0	1	1	0	0	0	0	0	0	34
CON	26	8	8	0	3	0	0	0	0	0	0	1	46
MIR	11	1	0	0	0	0	0	0	0	0	0	0	15
Totals	103	55	38	1	6	3	4	1	1	1	1	1	218

Table 3 b. Frequencies of individual heteroplasmic restriction enzyme mtDNA genotypes for each river.

Dde I polyacrylamide genotypes

River	BC	BE	BG	BGF	DE	EF	EH	Totals
STJ	0	0	0	0	2	0	0	2
ALT	0	0	0	0	0	0	0	0
SAN	0	0	0	0	0	0	0	0
CHO	0	0	0	0	0	0	0	0
RAP	0	0	1	0	1	0	0	2
DEL	2	0	0	0	0	0	0	2
HUD	0	0	1	0	0	0	0	1
CON	0	1	1	0	0	0	1	3
MR	0	0	0	1	0	1	0	2
Totals	2	1	3	1	3	1	1	12

Rsa I genotypes

River	AB	AC	BC	BD	Totals
STJ	1	0	4	1	6
ALT	0	0	0	0	0
SAN	0	0	0	0	0
CHO	0	0	0	0	0
RAP	0	0	0	0	0
DEL	1	2	0	0	3
HUD	1	1	0	0	2
CON	3	0	0	0	3
MR	0	0	0	0	0
Totals	6	3	4	1	14

Table 4. Kolmogorov-Smirnov (K-S) two-sample test values for determination of significance between rivers of genotypic frequency distributions.

	STJ	ALT	SAN	CHO	RAP	DEL	HUD	CON	MIR
STJ	X	0.825	1.259	*1.544	0.781	0.589	1.148	0.826	0.544
ALT	0.720	X	0.911	0.726	0.537	0.921	1.148	0.754	1.043
SAN	*1.567	0.669	X	1.021	0.320	0.802	0.369	0.633	0.630
CHO	0.242	0.834	*1.591	X	1.242	*1.448	*1.563	1.023	1.268
RAP	0.617	0.081	0.742	0.732	X	0.642	0.678	0.508	0.577
DEL	1.201	0.287	0.449	1.255	0.372	X	0.611	0.624	0.245
HUD	1.325	0.342	0.426	*1.365	0.428	0.293	X	0.690	0.432
CON	**2.45	1.165	0.417	**2.31	1.235	1.006	1.024	X	0.522
MIR	1.265	0.742	0.620	*1.324	0.730	0.585	0.856	0.807	X

above diagonal: K-S test Z value for Ava II genotypes
below diagonal: K-S test Z value for Dde I agarose genotypes

	STJ	ALT	SAN	CHO	RAP	DEL	HUD	CON	MIR
STJ	X	0.631	0.648	0.313	0.461	0.414	0.818	0.211	*1.617
ALT	0.094	X	0.214	0.421	0.502	0.363	0.503	0.576	*1.569
SAN	0.353	0.249	X	0.430	0.512	0.372	0.687	0.593	*1.598
CHO	0.389	0.217	0.324	X	0.373	0.275	0.621	0.421	*1.587
RAP	0.036	0.020	0.281	0.270	X	0.387	1.009	0.506	1.313
DEL	1.101	0.817	1.110	*1.357	0.900	X	0.682	0.313	*1.545
HUD	0.249	0.221	0.362	0.591	0.226	0.621	X	0.735	**1.90
CON	0.003	0.097	0.360	0.397	0.039	1.123	0.312	X	*1.517
MIR	0.517	0.441	0.593	0.781	0.464	0.359	0.522	0.607	X

above diagonal: K-S test Z value for Dde I polyacrylamide genotypes
below diagonal: K-S test Z value for Hinf I agarose genotypes

* $p \leq .05$
** $p < .01$

Table 4. Kolmogorov-Smirnov (K-S) two-sample test values for determination of significance between rivers of genotypic frequency distributions.

	STJ	ALT	SAN	CHO	RAP	DEL	HUD	CON	MIR
STJ	X	0.079	0.510	*1.644	0.000	1.064	0.612	0.022	0.000
ALT	**1.635	X	0.456	1.279	0.000	0.828	0.496	0.066	0.000
SAN	**1.850	1.193	X	0.919	0.490	0.387	0.017	0.542	0.490
CHO	**2.190	1.141	0.205	X	*1.409	0.618	1.104	*1.718	*1.409
RAP	**1.879	0.272	*1.384	*1.332	X	0.920	0.555	0.075	0.000
DEL	**1.920	0.456	1.093	1.033	0.462	X	0.487	1.126	0.920
HUD	**2.183	0.219	*1.357	1.305	0.304	0.384	X	0.662	0.555
CON	**1.979	0.313	1.234	1.117	0.488	0.395	0.373	X	0.075
MIR	**1.986	0.516	*1.462	*1.415	0.346	0.531	0.645	0.600	X

above diagonal: K-S test Z value for Hinf I poly acrylamide genotypes
below diagonal: K-S test Z value for Rsa I genotypes
* p <.05
** p <.01

Table 5. Derivation of S values to test significance of Kendall's rank- correlation coefficient.

River	Freq (%) of genotype Ava II A	P	Q	P-Q	Freq (%) of genotype Ava II B	P	Q	P-Q
FLA	32	8	0	8	44	2	6	-4
ALT	38	7	0	7	25	4	3	1
SAN	68	0	6	-6	21	4	1	3
CHO	43	4	1	3	14	5	0	5
RAP	57	1	3	-2	21	4	0	4
DEL	44	3	0	3	48	0	3	-3
HUD	60	0	2	-2	34	1	1	0
CON	51	0	1	-1	33	1	0	1
MIR	47	-	-	-	47	-	-	-
				S = 10				S = 7

River	Freq (%) of genotype Dde I Ag A	P	Q	P-Q	Freq (%) of genotype Dde I Ag B	P	Q	P-Q
FLA	34	7	1	6	63	0	8	-8
ALT	56	5	2	3	44	2	5	-3
SAN	79	1	5	-4	16	4	2	2
CHO	29	5	0	5	62	0	5	-5
RAP	53	4	0	4	47	0	4	-4
DEL	65	3	0	3	27	1	2	-1
HUD	67	2	0	2	33	0	2	-2
CON	90	0	1	-1	7	0	1	-1
MIR	73	-	-	-	0	-	-	-
				S = 18				S = -22

River	Freq (%) of genotype Hinf I Ag A	P	Q	P-Q	Freq (%) of genotype Hinf I Ag B	P	Q	P-Q
FLA	76	4	3	1	24	4	4	0
GA	78	2	5	-3	22	5	2	3
SC	80	1	5	-4	10	6	0	6
NC	86	0	5	-5	14	5	0	5
VA	77	0	4	-4	23	4	0	4
DEL	45	3	0	3	55	0	3	-3
HUD	69	1	1	0	31	1	1	0
CON	76	0	1	-1	24	1	0	1
MIR	58	-	-	-	42	-	-	-
				S = 13				S = 16

Table 5. Derivation of S values to test significance of Kendall's rank- correlation coefficient.

River	Freq (%) of genotype Rsa I A	P	Q	P-Q	Freq (%) of genotype Rsa I B	P	Q	P-Q
FLA	12	8	0	8	18	5	3	2
GA	67	2	5	-3	8	6	1	5
SC	22	6	0	6	61	1	5	-4
NC	25	5	0	5	65	0	5	-5
VA	71	1	3	-2	14	3	1	2
DEL	56	3	0	3	24	0	3	-3
HUD	62	1	1	0	21	0	2	-2
CON	57	1	0	1	17	0	1	-1
MIR	73	-	-	-	7	-	-	-
				S = 18				S = -6

River	Freq (%) of genotype Rsa I C	P	Q	P-Q
FLA	50	0	8	-8
GA	25	0	7	-7
SC	11	2	4	-2
NC	10	2	3	-1
VA	7	2	2	0
DEL	4	2	1	1
HUD	12	1	1	0
CON	17	0	1	-1
MIR	0	-	-	-
				S = -18

Table 6. Kendall's rank-correlation coefficient and significance.

Enzyme	Ava II	Ava II	Dde I ag	Dde I ag	Hinf I ag	Hinf I ag	Hinf I ac	Hinf I ac	Rsa I	Rsa I	Rsa I
Genotype	A	B	A	B	A	B	A	B	A	B	C
S	10	7	18	-22	13	16	5	-5	18	-6	-18
K	0.28	0.19	0.5	*-0.61	0	0.44	0.14	-0.1	0.5	0.17	0.5

*p<.05

Table 7. Composite mtDNA genotypes for American shad from nine spawning rivers.

Composite number	Composite*	STJ	ALT	SAN	CHO	RAP	DEL	HUD	CON	MIR	Totals
1	AAAAAAA	0	0	0	1	0	2	3	5	1	12
2	AAAAAAB	1	0	1	1	0	1	1	1	0	6
3	AAAAAAC	0	0	0	0	0	1	0	1	0	2
4	AAAAAAH	0	0	0	0	0	0	0	1	0	1
5	AAAABAA	0	0	0	0	0	1	3	1	0	5
6	AAAABA(AB)	0	0	0	0	0	0	1	0	0	1
7	AAAABAB	0	0	0	0	0	1	0	2	0	3
8	AAAACAB	0	0	1	0	0	0	0	0	0	1
9	AABAAAA	0	0	0	0	0	0	1	0	0	1
10	AA(BG)AAA	0	0	0	0	0	0	0	1	0	1
11	AA(BG)AAAE	0	0	0	0	1	0	0	0	0	1
12	AA(BG)ABAB	0	0	0	0	0	0	1	0	0	1
13	AACAAAA	0	0	1	0	0	0	0	0	0	1
14	AACAAAC	0	0	0	1	0	0	0	0	0	1
15	AACABAB	0	0	2	0	0	0	0	0	0	2
16	AADAAAA	0	0	0	0	1	0	0	0	0	1
17	AADAAAC	0	1	0	0	0	0	0	0	0	1
18	AADAAAB	0	0	1	0	0	0	0	0	0	1
19	AAEAAAB	0	0	0	0	1	0	0	0	0	1
20	AA(EF)ABAA	0	0	0	0	0	0	0	0	1	1
21	AA(EH)AAAA	0	0	0	0	0	0	0	1	0	1
22	AAFAAAE	0	0	0	0	0	0	0	0	2	2
23	AAGAB(AB)	0	0	0	0	0	0	0	1	0	1
24	AAKAAAA	0	0	0	0	0	0	0	1	0	1
25	ABAAAAA	1	2	0	0	3	0	1	0	0	7
26	ABAAAAB	1	0	0	2	1	0	0	0	0	4
27	ABAAABA	0	0	1	0	0	1	0	0	0	2
28	ABAAAB(AB)	0	0	0	0	0	1	0	0	0	1
29	ABAAAAC	0	0	0	1	0	0	0	0	0	1
30	ABAAABB	0	0	0	1	0	1	2	0	0	4
31	ABAABAA	0	0	0	0	1	0	0	0	0	1
32	ABAAB(AC)	0	0	0	0	0	1	0	0	0	1
33	ABAABAC	0	0	0	0	0	0	0	1	0	1
34	ABABAAB	1	0	0	0	0	0	0	0	0	1
35	ABABAA(BD)	1	0	0	0	0	0	0	0	0	1
36	ABABAAC	1	0	0	0	0	0	0	0	0	1
37	ABABAAA	0	0	0	0	0	0	1	0	0	1
38	ABABBAA	1	0	0	0	0	0	0	0	0	1
39	AB(BC)AAAA	0	0	0	0	0	1	0	0	0	1
40	ABEAAAB	1	0	0	0	0	0	0	0	0	1
41	ACCAAAB	0	0	0	0	0	0	0	1	1	2
42	ADAABAD	1	0	0	0	0	0	0	0	0	1
43	AEAAAAA	0	0	1	0	0	0	0	0	0	1
44	BAAAAAA	0	1	0	1	1	0	3	4	1	11
45	BAAAAA(AB)	0	0	0	0	0	0	0	1	0	1
46	BAAAAAB	0	0	2	0	0	0	0	1	0	3
47	BAAAAB(BC)	1	0	0	0	0	0	0	0	0	1
48	BAAABAA	0	0	0	0	0	3	1	0	1	5
49	BAAABAB	0	0	0	0	0	0	0	2	0	2
50	BAAABAC	0	0	0	0	0	0	1	0	0	1
51	BAABBA(AB)	1	0	0	0	0	0	0	0	0	1
52	BA(BE)ABAA	0	0	0	0	0	0	0	1	0	1
53	BA(BGF)AAAA	0	0	0	0	0	0	0	0	1	1
54	BACAAAC	1	0	0	0	0	0	0	0	0	1
55	BACABAA	0	1	0	0	1	0	0	0	0	1

Table 7. Composite mtDNA genotypes for American shad from nine spawning rivers.

Composite number	Composite*	STJ	ALT	SAN	CHO	RAP	DEL	HUD	CON	MIR	Totals
56	BACABAB	0	0	0	0	0	1	0	0	0	1
57	BA(DE)AAA(BC)	1	0	0	0	0	0	0	0	0	1
58	BAEABAC	0	0	0	0	0	0	0	1	0	1
59	BAFABAA	0	0	0	0	0	0	0	0	1	1
60	BBAAAAA	0	0	0	0	0	0	1	0	0	1
61	BBAAAAC	4	0	0	0	0	0	0	0	0	4
62	BBAABAB	0	0	0	0	0	0	1	0	0	1
63	BBAABBA	0	0	0	1	0	0	0	0	0	1
64	BBAABBF	0	0	0	0	0	1	0	0	0	1
65	BBABAA(BC)	1	0	0	0	0	0	0	0	0	1
66	BBABAAC	1	0	0	0	0	0	0	0	0	1
67	BBBABAA	0	0	0	0	0	0	1	0	0	1
68	BBEABAC	1	0	0	0	0	0	0	0	0	1
69	BBEBAAC	1	0	0	0	0	0	0	0	0	1
70	BBGABA(BC)	1	0	0	0	0	0	0	0	0	1
71	BCCABAB	0	0	0	1	0	0	0	0	0	1
72	BC(CF)ABAE	0	0	0	0	0	0	0	0	1	1
73	CADAAAC	1	0	0	0	0	0	0	0	0	1
74	CA(DE)AAAB	1	0	0	0	0	0	0	0	0	1
75	CAEAAAA	1	0	0	0	0	0	0	0	0	1
76	EBAAABA	0	0	0	0	0	0	1	0	0	1
77	FADAAAC	0	1	0	0	0	0	0	0	0	1
78	FBAAAAA	1	0	0	0	0	0	0	0	0	1
79	GABAAAC	0	0	0	0	0	0	1	0	0	1
80	GBAAAAC	1	0	0	0	0	0	0	0	0	1
81	GBAABAC	1	0	0	0	0	0	0	0	0	1
82	HCCAAAA	0	0	0	0	0	1	0	0	0	1
83	IAAABAC	1	0	0	0	0	0	0	0	0	1
84	JAAAAAC	1	0	0	0	0	0	0	0	0	1
85	KA(DE)ABAC	0	0	0	0	1	0	0	0	0	1
86	KBAAAAA	0	0	0	0	1	0	0	0	0	1
87	KBAABAC	0	1	0	0	0	0	0	0	0	1
88	LAAAAAA	0	0	0	0	0	0	0	2	0	2
89	LAAAAA(AB)	0	0	0	0	0	0	0	1	0	1
90	LAAAAAC	0	0	0	0	0	0	0	1	0	1
91	LAIAAAC	0	0	0	0	0	0	0	1	0	1
92	LBAAABA	0	0	0	1	0	0	0	0	0	1
93	LBAAABB	0	0	1	2	0	0	0	0	0	3
94	LBHAAAB	0	0	0	1	0	0	0	0	0	1
95	LFDABBB	0	0	0	1	0	0	0	0	0	1
96	MBAAAAC	0	0	0	0	0	0	0	1	0	1
97	MBEAABB	0	0	0	1	0	0	0	0	0	1
98	NBAAAAB	0	0	0	1	0	0	0	0	0	1
Total number of samples		30	7	11	17	12	17	24	33	10	161
Total number of genotypes within river		27	6	9	15	10	15	17	23	9	
Total number of genotypes unigue to that specific river		24	3	5	9	6	6	10	15	5	83

*Letters from left to right indicate digestion patterns for Ava II agarose and polyacrylamide gels, Dde I agarose gels, Dde I polyacrylamide gels, EcoR I agarose gels, Hinf I agarose gels, Hinf I polyacrylamide gels, and Rsa I polyacrylamide gels.

Table 8. Composite mtDNA genotypes shared among rivers.

Composite* number	Composite**	STJ	ALT	SAN	CHO	RAP	DEL	HUD	CON	MIR	TOTALS
1 (1)	AAAAAAA	0	0	0	1	0	2	3	5	1	12
2 (2)	AAAAAAB	1	0	1	1	0	1	1	1	0	6
3 (3)	AAAAAAC	0	0	0	0	0	1	0	1	0	2
5 (4)	AAAABAA	0	0	0	0	0	1	3	1	0	5
7 (5)	AAAABAB	0	0	0	0	0	1	0	2	0	3
25 (6)	ABAAAAA	1	2	0	0	3	0	1	0	0	7
26 (7)	ABAAAAB	1	0	0	2	1	0	0	0	0	4
27 (8)	ABAAABA	0	0	1	0	0	1	0	0	0	2
30 (9)	ABAAABB	0	0	0	1	0	1	2	0	0	4
41 (10)	ACCAAAB	0	0	0	0	0	0	0	1	1	2
44 (11)	BAAAAAA	0	1	0	1	1	0	3	4	1	11
46 (12)	BAAAAAB	0	0	2	0	0	0	0	1	0	3
48 (13)	BAAABAA	0	0	0	0	0	3	1	0	1	5
55 (14)	BACABAA	0	1	0	0	1	0	0	0	0	2
93 (15)	LBAAABB	0	0	1	2	0	0	0	0	0	3
Total number of samples		3	4	5	8	6	11	14	16	4	71

**Letters from left to right indicate digestion patterns for Ava II agarose and acrylamide gels, Dde I agarose gels, Dde I acrylamide gels, EcoR I agarose gels, Hinf I agarose gels, Hinf I acrylamide gels, and Rsa I acrylamide gels.

Numbers on te left are the original (of 98) composite mtDNA genotypes.
Numbers in parentheses are the new numbers for the mtDNA composite genotypes that are shared among rivers.

TABLE 9. UPHOLT'S F AND NEI'S P VALUES FOR SHARED COMPOSITE AND HICKORY SHAD MTDNA GENOPTYPES.

	1 AAAAAAA	2 AAAAAAB	3 AAAAAAC	4 AAAABAA	5 AAAABAB	6 ABAAAAA	7 ABAAAAB	8 ABAAABA	9 ABAAABB	10 ACCAAAB	11 BAAAAAA	12 BAAAAAB	13 BAAABAA	14 BACABAA	15 LBAAABB	hickory shad
1 AAAAAAA	X	0.993	0.993	0.991	0.984	0.983	0.978	0.974	0.967	0.966	0.983	0.978	0.974	0.985	0.959	0.317
2 AAAAAAB	0.00057	X	0.993	0.984	0.991	0.976	0.983	0.967	0.974	0.973	0.976	0.983	0.987	0.958	0.966	0.317
3 AAAAAAC	0.00118	0.00118	X	0.984	0.977	0.976	0.969	0.967	0.96	0.958	0.976	0.969	0.987	0.958	0.952	0.317
4 AAAABAA	0.0006	0.00117	0.0017	X	0.993	0.983	0.967	0.966	0.959	0.957	0.974	0.967	0.983	0.974	0.951	0.325
5 AAAABAB	0.00117	0.0006	0.00178	0.00057	X	0.967	0.983	0.966	0.966	0.964	0.967	0.974	0.976	0.967	0.958	0.325
6 ABAAAAA	0.00115	0.00172	0.00172	0.00175	0.00232	X	0.993	0.984	0.984	0.966	0.966	0.959	0.957	0.957	0.976	0.317
7 ABAAAAB	0.00172	0.00115	0.00293	0.00232	0.00175	0.00057	X	0.984	0.991	0.973	0.959	0.966	0.95	0.95	0.983	0.317
8 ABAAABA	0.00175	0.00232	0.00232	0.0023	0.00287	0.0006	0.00117	X	0.993	0.957	0.957	0.95	0.949	0.949	0.985	0.317
9 ABAAABB	0.00232	0.00175	0.00293	0.00287	0.0023	0.00117	0.0006	0.00057	X	0.964	0.95	0.957	0.942	0.942	0.992	0.317
10 ACCAAAB	0.00242	0.00185	0.00303	0.00302	0.00245	0.00242	0.00185	0.00302	0.00245	X	0.949	0.956	0.94	0.949	0.956	0.31
11 BAAAAAA	0.0012	0.00177	0.00177	0.0018	0.00237	0.00235	0.00292	0.00295	0.00352	0.00362	X	0.993	0.991	0.982	0.959	0.317
12 BAAAAAB	0.00177	0.0012	0.003	0.00237	0.0018	0.00292	0.00235	0.00352	0.00295	0.00305	0.00177	X	0.984	0.975	0.966	0.317
13 BAAABAA	0.0018	0.00237	0.00237	0.0012	0.00177	0.00295	0.00352	0.0035	0.00407	0.00422	0.0006	0.00117	X	0.991	0.951	0.325
14 BACABAA	0.00241	0.00238	0.00238	0.00181	0.00238	0.00298	0.00355	0.00353	0.0041	0.00422	0.00121	0.00178	0.00061	X	0.951	0.317
15 LBAAABB	0.00269	0.00232	0.0035	0.00344	0.00287	0.00174	0.00117	0.00174	0.00057	0.00302	0.00233	0.00233	0.00345	0.0024	X	0.317
hickory shad	0.087	0.087	0.087	0.085	0.085	0.087	0.087	0.087	0.087	0.089	0.087	0.087	0.085	0.087	0.087	X

F VALUES ABOVE THE DIAGONAL
P VALUES BELOW THE DIAGONAL

Table 10. Autapomorphic and synapomorphic fragments for each single-enzyme genotype.

Ava II fragments that vary	Genotypes													
Number	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1													2857	
2	2371	2395	2375	2390	2390	2390	2404	2536	2366	2385	2385	2457		2437
3	2134	2134	2158	2180	2180	2150	2163		2148	2171	2158	2211	2223	2184
4				1900				1940						
5	1870	1867	1881	1800		1870	1881		1839	1867	1881	1872	1869	1862
6		1710	1730		1690		1730			1737	1737			
7	1677			1630		1648						1642	1655	1641
8			1578			1515	1507	1560	1555	1532				1584
9									1486				1497	
10	1462		1445								1477	1455		
11		1425		1430	1400	1433		1405	1407	1433				1438
12	1190	1194		1180	1200	1216	1218	1279	1194	1209	1216	1248	1243	
13					1100									
14	958	952		955	920	962	962	954		956	964	890	890	934
15					715									
16				583										
17			557											
18	351	351		354	351	351	351	?	?	?	?	?	?	?
19			325					?	?	?	?	?	?	?
20	287	287	280		287	287	287	?	?	?	?	?	?	?

Dde I ag fragments that vary	Genotypes					
Number	A	B	C	D	E	F
1			1430		1500	
2						1260
3		1015		1015	1015	1015
4	735	735	735		735	735
5					695	
6	650	650		650		650
7	595	595		595	595	595
8	540					

Table 10. Autapomorphic and synapomorphic fragments for each single-enzyme genotype.

Dde I ac fragments that vary Genotypes

Number	A	B	C	D	E	F	G	H	I	J	K	BC	BE	BG	BGF	DE	EF	EH
1														429				
2							404								402			
3	394	394	395	394	394	388	295	385	398			395	398	391	388	394	393	391
4		377							362			375	376	356	364			
5					335								335			334	328	335
6	327	331	326	322	327	325	326		325	326	326	330	326	328	325	327	325	
7	315	319	317	312	315	316	315	315		316	315	320	318	317	316	316	315	317
8									289									
9	271	275		253	271	261	267	267	268	260	267		269	253	261	253	261	268
10	271	275		253	271	261	267	267	268	260	267		269	253	261	253	261	268
11	245	251	238	238	245		242	242	242	238	242	235	246	238		238		243
12	169	175	166	168	169	169		161	166	165		165	170	168	170	165	169	169
13	156		153	157	156	156	153	153	158	155	153	155	149	157	154		153	156

EcoR I fragments that vary Genotypes

Number	A	B
1	9973	
2		5097
3		4220

Hinf I ag fragments that vary Genotypes

Numbe	A	B	C	D
1		1065		
2	1040			
3	1011	1000		
4				840

Hinf I ac fragments that vary Genotypes

Numbe	A	B
1		225
2	175	

Rsa I fragments that vary Genotypes

Number	A	B	AB	C	AC	BC	D	BD	E	F	G	H
1		365	365			365		365				
2	355		355	355	355	355			355	355	343	
3											338	342
4				330	330	330						
5	325	328	325		325						322	323
6									315		314	
7										145	?	?
8										105	?	?
9	122	122	122	122	122	122	122	122			?	?
10	90	90	90	90	90	90	90	90	90		?	?

Table 11. HENNIG86 coding of gragments for each single enzyme mtDNA genotype.

Ava II

Genotype	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	0	1	1	0	1	0	1	0	0	1	0	1	0	1	0	0	0	1	0	1
B	0	1	1	0	1	1	0	0	0	0	1	1	0	1	0	0	0	1	0	1
C	0	1	1	0	1	1	0	1	0	1	0	0	0	0	0	0	1	0	1	1
D	0	1	1	1	1	0	1	0	0	0	1	1	0	1	0	1	0	1	0	0
E	0	1	1	0	0	1	0	0	0	0	1	1	1	1	1	0	0	1	0	1
F	0	1	1	0	1	0	1	1	0	0	1	1	0	1	0	0	0	1	0	1
G	0	1	1	0	1	1	0	1	0	0	0	1	0	1	0	0	0	1	0	1
H	0	1	0	1	0	0	0	1	0	0	1	1	0	1	0	0	0	?	?	?
I	0	1	1	0	1	0	0	1	1	0	1	1	0	0	0	0	0	?	?	?
J	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0	0	0	?	?	?
K	0	1	1	0	1	1	0	0	0	1	0	1	0	1	0	0	0	?	?	?
L	0	1	1	0	1	0	1	0	0	1	0	1	0	1	0	0	0	?	?	?
M	1	0	1	0	1	0	1	0	1	0	0	1	0	1	0	0	0	?	?	?
N	0	1	1	0	1	0	1	1	0	0	1	0	0	1	0	0	0	?	?	?
hickory shad	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Dde I

Genotype agarose	1	2	3	4	5	6	7	8
A	0	0	0	1	0	1	1	1
B	0	0	1	1	0	1	1	0
C	1	0	0	1	0	0	0	0
D	0	0	1	0	0	1	1	0
E	1	0	1	1	1	0	1	0
F	0	1	1	1	0	1	1	0
hickory shad	?	?	?	?	?	?	?	?

Dde I

Genotype Acrylamide	1	2	3	4	5	6	7	8	9	0	11	12	13
A	0	0	1	0	0	1	1	0	1	1	1	1	1
B	0	0	1	1	0	1	1	0	1	1	1	1	0
C	0	0	1	0	0	1	1	0	0	0	1	1	1
D	0	0	1	0	0	1	1	0	1	1	1	1	1
E	0	0	1	0	1	1	1	0	1	1	1	1	1
F	0	0	1	0	0	1	1	0	1	1	0	1	1
G	0	1	1	0	0	1	1	0	1	1	1	0	1
H	0	0	1	0	0	0	1	0	1	1	1	1	1
I	0	0	1	1	0	1	0	1	1	1	1	1	1
J	0	0	0	0	0	1	1	0	1	1	1	1	1
K	0	0	0	0	0	1	1	0	1	1	1	0	1
BC	0	0	1	1	0	1	1	0	0	0	1	1	1
BE	0	0	1	1	1	1	1	0	1	1	1	1	1
BG	1	0	1	1	0	1	1	0	1	1	1	1	1
BGF	0	1	1	1	0	1	1	0	1	1	0	1	1
CF	0	0	1	0	0	1	1	0	0	0	0	1	1
DE	0	0	1	0	0	1	1	0	1	1	1	1	0
EF	0	0	1	0	0	1	1	0	1	1	0	1	1
EH	0	0	1	0	0	0	1	0	1	1	1	1	1
hickory shad	0	0	0	0	0	0	0	0	1	1	0	1	0

Table 11. HENNIG86 coding of gragments for each single enzyme mtDNA genotype.

EcoR I
Genotype

	1	2	3
A	1	0	0
B	0	1	1
hickory shad	1	0	1

Hinf I
agarose

genotype	1	2	3	4
A	0	1	1	0
B	1	0	0	0
C	1	1	0	1
D	0	0	0	1
hickory shad	1	0	0	0

Hinf I
Acrylamide

genotype	1	2
A	0	1
B	1	0
hickory shad	0	1

Rsa I

genotype	1	2	3	4	5	6	7	8	9	10
A	0	1	0	0	1	0	0	0	1	1
B	1	0	0	0	1	0	0	0	1	1
AB	1	1	0	0	1	0	0	0	1	1
C	0	1	0	1	0	0	0	0	1	1
AC	0	1	0	1	1	0	0	0	1	1
BC	1	1	0	1	0	0	0	0	1	1
D	0	0	0	0	0	0	0	0	1	1
BD	1	0	0	0	0	0	0	0	1	1
E	0	1	0	0	0	1	0	0	0	1
F	0	1	0	0	0	0	1	1	0	0
G	0	1	1	0	1	1	?	?	?	?
H	0	0	1	0	1	0	?	?	?	?
hickory shad	0	0	0	0	0	0	0	0	0	0

Table 12. HENNIG86 code for shared composite and hickory shad mtDNA genotypes.

Composite number		AvaII	DdeI ag	DdeI ac
1	AAAAAAA	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
2	AAAAAAB	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
3	AAAAAAC	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
5	AAAABAA	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
7	AAAABAB	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
25	ABAAAAA	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 1 1 0 1 1 0	0 0 1 0 0 1 1 0 1 1 1 1 1
26	ABAAAAB	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 1 1 0 1 1 0	0 0 1 0 0 1 1 0 1 1 1 1 1
27	ABAAABA	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 1 1 0 1 1 0	0 0 1 0 0 1 1 0 1 1 1 1 1
30	ABAAABB	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	0 0 1 1 0 1 1 0	0 0 1 0 0 1 1 0 1 1 1 1 1
41	ACCAAAB	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 1 0 1	1 0 0 1 0 0 0 0	0 0 1 0 0 1 1 0 0 0 1 1 1
44	BAAAAAA	0 1 1 0 1 1 0 0 0 0 1 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
46	BACABAA	0 1 1 0 1 1 0 0 0 0 1 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 0 0 1 1 1
48	BAAABAA	0 1 1 0 1 1 0 0 0 0 1 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 1 1 1 1 1
55	BACABAA	0 1 1 0 1 1 0 0 0 0 1 1 0 1 0 0 0 1 0 1	0 0 0 1 0 1 1 1	0 0 1 0 0 1 1 0 0 0 1 1 1
93	LBAAABB	0 1 1 0 1 0 1 0 0 1 0 1 0 1 0 0 0 ? ? ?	0 0 1 1 0 1 1 0	0 0 1 0 0 1 1 0 1 1 1 1 1
hickory shad		0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	? ? ? ? ? ? ? ?	0 0 0 0 0 0 0 0 1 1 0 1 0

Composite number		EcoRI	HinfI ag	HinfI ag	RsaI
1	AAAAAAA	1 0 0	0 1 1 0	0 1	0 1 0 0 1 0 0 0 1 1
2	AAAAAAB	1 0 0	0 1 1 0	0 1	1 0 0 0 1 0 0 0 1 1
3	AAAAAAC	1 0 0	0 1 1 0	0 1	0 1 0 1 0 0 0 0 1 1
5	AAAABAA	1 0 0	1 0 1 0	0 1	0 1 0 0 1 0 0 0 1 1
7	AAAABAB	1 0 0	1 0 1 0	0 1	1 0 0 0 1 0 0 0 1 1
25	ABAAAAA	1 0 0	0 1 1 0	0 1	0 1 0 0 1 0 0 0 1 1
26	ABAAAAB	1 0 0	0 1 1 0	0 1	1 0 0 0 1 0 0 0 1 1
27	ABAAABA	1 0 0	0 1 1 0	1 0	0 1 0 0 1 0 0 0 1 1
30	ABAAABB	1 0 0	0 1 1 0	1 0	1 0 0 0 1 0 0 0 1 1
41	ACCAAAB	1 0 0	0 1 1 0	0 1	1 0 0 0 1 0 0 0 1 1
44	BAAAAAA	1 0 0	0 1 1 0	0 1	0 1 0 0 1 0 0 0 1 1
46	BACABAA	1 0 0	1 0 1 0	0 1	0 1 0 0 1 0 0 0 1 1
48	BAAABAA	1 0 0	1 0 1 0	0 1	0 1 0 0 1 0 0 0 1 1
55	BACABAA	1 0 0	1 0 1 0	0 1	0 1 0 0 1 0 0 0 1 1
93	LBAAABB	1 0 0	0 1 1 0	1 0	1 0 0 0 1 0 0 0 1 1
hickory shad		1 0 1	1 0 0 0	0 1	0 0 0 0 0 0 0 0 0 0

Table 13. Fragments shared between hickory shad mtDNA genotypes and "A" mtDNA genotypes of American shad.

Fragment sizes in bp:									
Ava II ag		Dde I ag		EcoR I		Hinf I ag & ac		Rsa I	
A. shad	H. shad	A. shad	H. shad	A. shad	H. shad	A. shad	H. shad	A. shad	H. shad
	4072	509	507	9973	9973		1254		653
	3054	496	499	7315	7315	1191			583
2371		479			4500		1100	551	
	2200		476	1377			1044		524
2134		468				940		519	
	2060		450				899	498	
1870		440				887			485
1718	1718	440					761		485
1677			438			704			450
1462			419			647		444	
1190		416					630	444	
958			396			621			408
864		386					575	398	397
529			373			571		355	
			357				570	326	
		344				557			318
			341				550	309	
		322				536		292	
		311	309			510			288
			298				506	282	
		293					502	275	274
		293					502	266	264
			291			500		244	241
		289				485			236
		281	281				480	220	221
		264	263			477	477	196	194
		264	263			477	477		190
			260			434		186	
		240					432		179
		221				430		165	
		198	200			357			162
		198	200			347		122	
		184	184				340		91
		171	173				320	89	
		164				313	313		
			159				298		
		157				296			
		152				268			
		149	149				260		
		132	132				245		
						235	235		
						218			
							200		
						193	193		
						188			
							180		
						178	178		
						175	175		
						171	171		
						136			
Total						118			
no. of frags 10	5	28	24	3	5	30	29	20	20
no. frags shared =	1		12		2		8		5
F value	0.133		0.461		0.333		0.271		0.25
p value	0.152		0.055		0.066		0.095		0.127

Fig. 1. A map of the Atlantic coast of North America depicting locations of the nine collection rivers of American shad. They are from south to north: A. St. Johns River, FLA; B. Altamaha River, GA; C. Santee-Cooper rivers, SC; D. Chowan and Albermarle Sound, NC; E. Rappahannock River, VA; F. Delaware River, NJ/NY/PA; G. Hudson River, NY; H. Connecticut River, MA; and I. Miramichi River, New Brunswick, Canada.

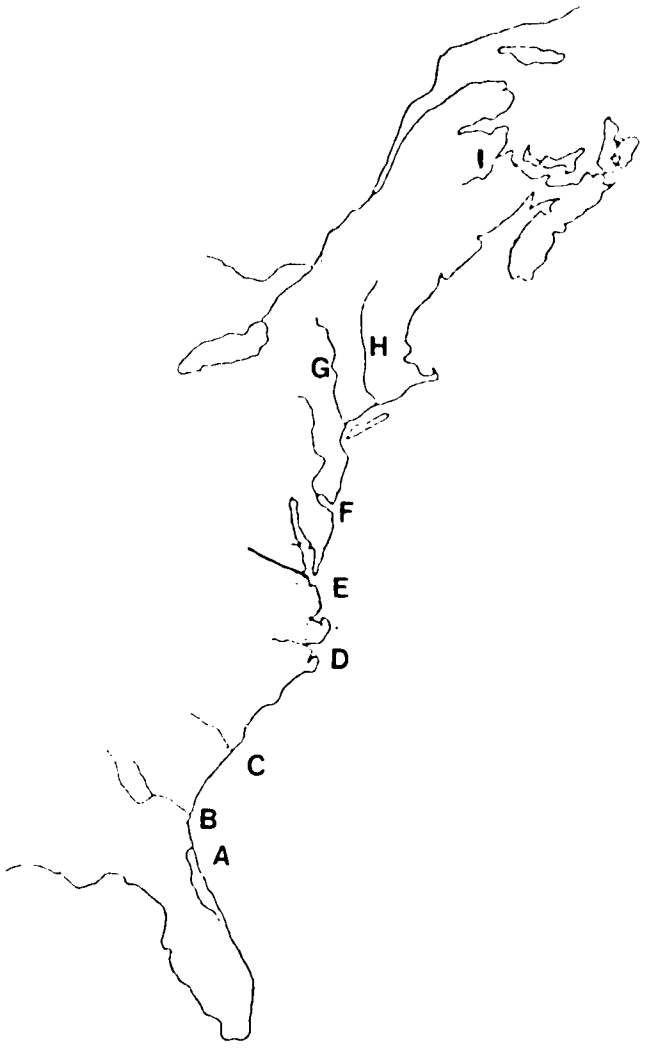

I
H
G
F
E
D
C
B
A

Fig. 2. An autoradiograph of Ava II 1.4 % agarose gel digestion patterns. Lanes A and C depict Genotype A, lane B is indicative of Genotype B, and lane D is lambda DNA cut with Hind III.

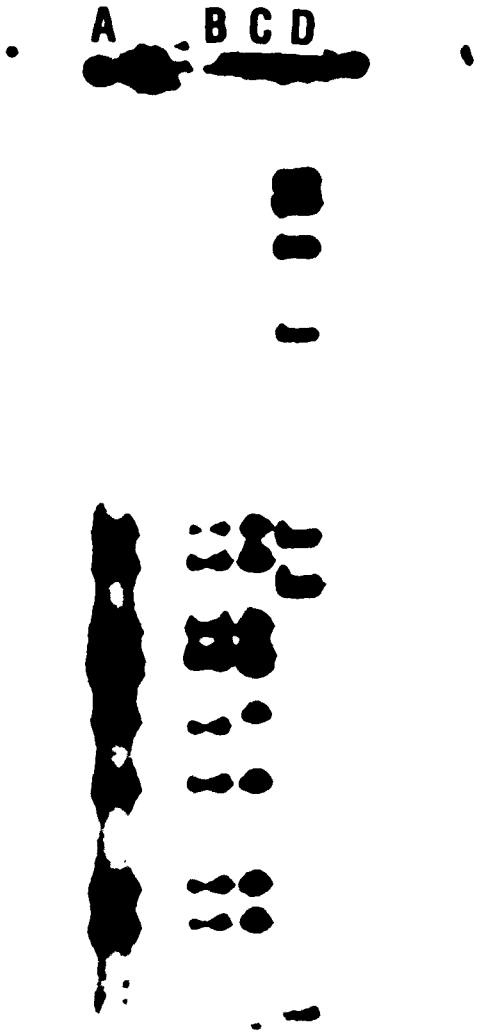

A
B C D

Fig. 3. An autoradiograph of Ava II digests separated on a 6% polyacrylamide gel. Lane C is representative of Genotype D, lane I depicts Genotype C, and lane L is the 1 kb molecular weight standard. The rest of the lanes are representative of Genotypes A and B.


A B C D E F G H I J K L
506
396
344
298
210
200

Fig. 4. An autoradiograph of Dde I digests of American shad mtDNA separated on a 1.6% agarose gel. Polymorphism can be seen in the presence/absence of 1015 bp and 540 bp fragments as indicated by the two arrows. Representatives of Genotype A can be seen in lanes b, h, i, k, and l while samples in lanes a, c, d, e, f, and j depict Genotype B. The molecular weight standard, a 1 kb ladder, is in lane g. Molecular sizes (bp) are indicated on left.

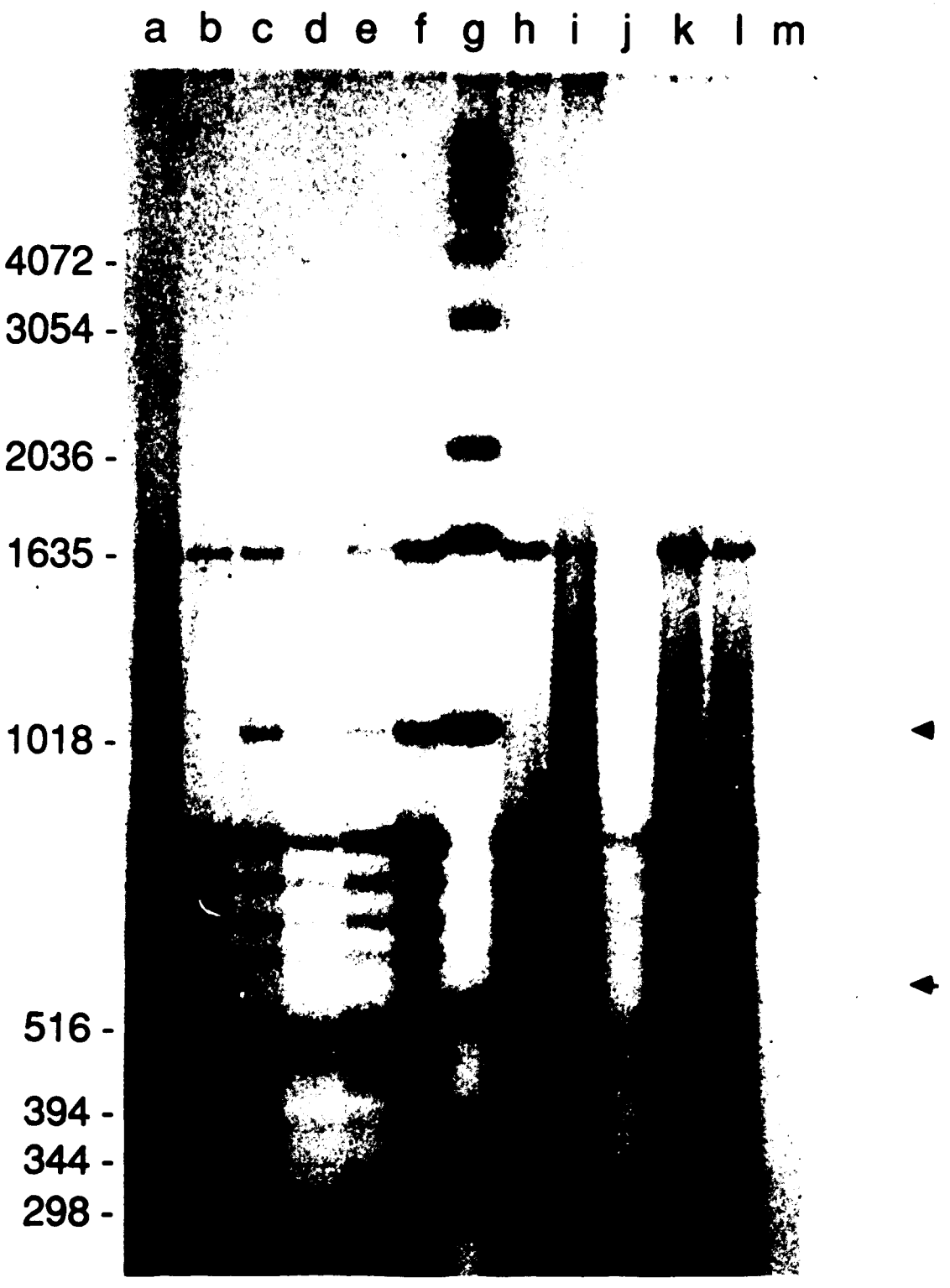

a b c d e f g h i j k l m
4072 -
3054 -
2036 -
1635 -
1018 -
516 -
394 -
344 -
298 -

Fig 5. An autoradiograph of Dde I digests fractionated on a 6% polyacrylamide gel. Six different mtDNA genotypes can be seen. Polymorphic regions are indicated by the arrows. Lane to genotype designations are as follows: a-Genotype A; b-Genotype BG; c-no sample; d-molecular weight standard, 1 kb ladder; e-Genotype BGF; f-Genotype F; g-Genotype EF; h-Genotype F; i-genotype F; j-Genotype A; k-Genotype C; l-Genotype A. Any genotype containing the letter "F" is indigenous to the Miramichi River.

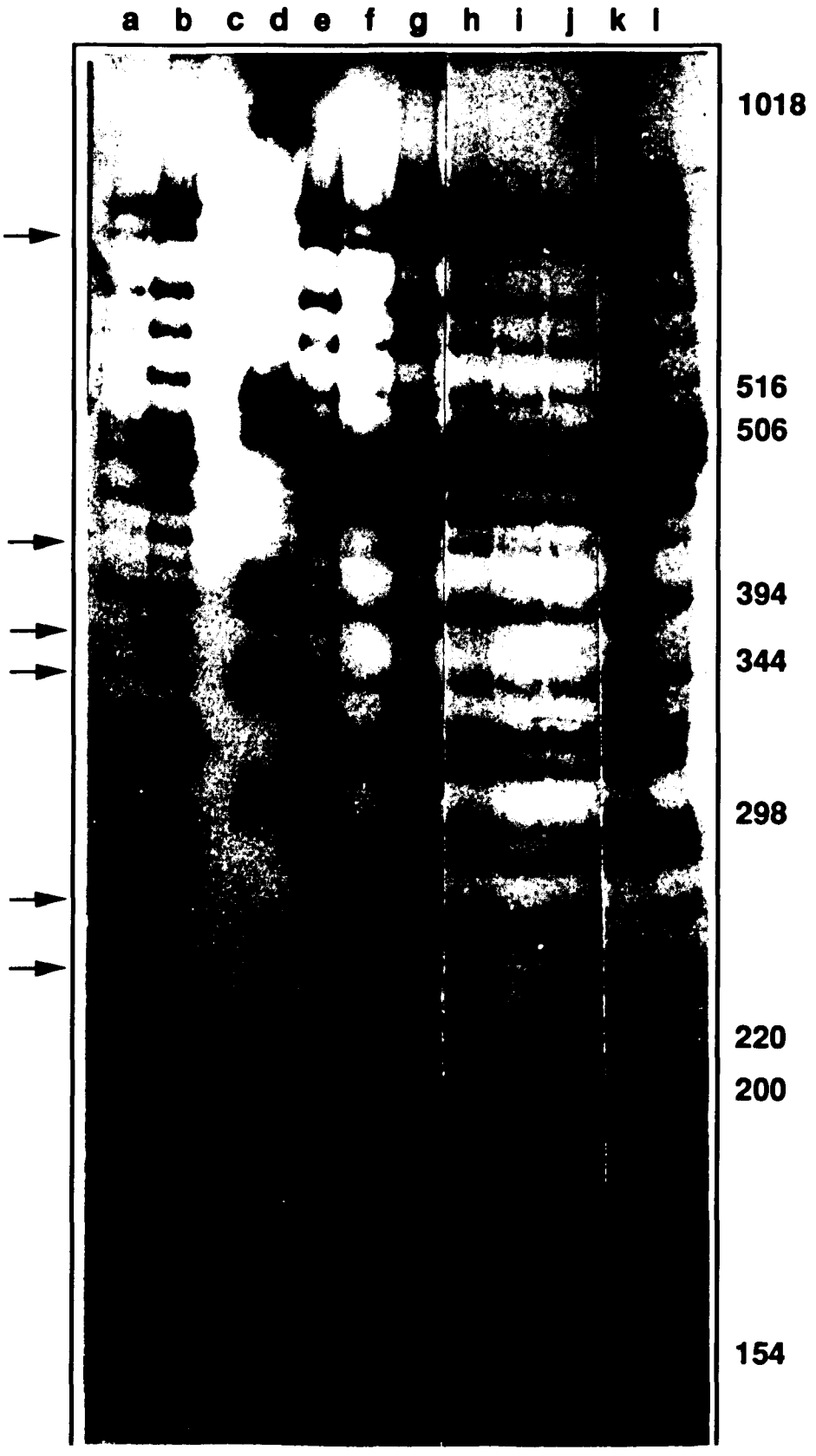

a b c d e f g h i j k l
1018
516
506
394
344
298
220
200
154

Fig. 6. Representative Hinf I digestion patterns separated on a 1.6% agarose gel. Lanes a, e, f, g, and j depict Genotype A while lanes i and m display Genotype B. Lane c is the molecular weight marker PBR322.

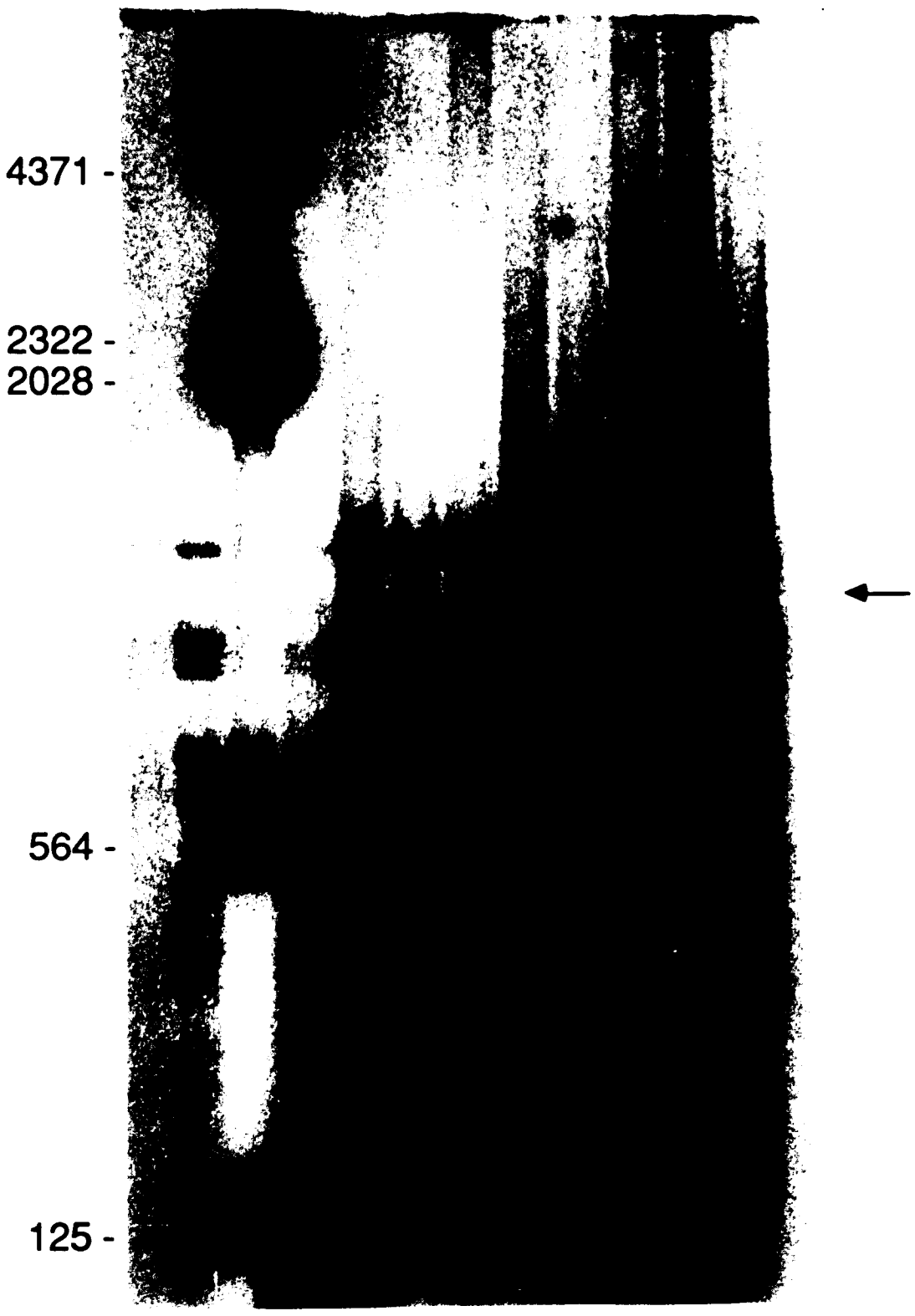

a b c d e f g h i j k l m
4371 -
2322 -
2028 -
564 -
125 -

Fig. 7. An autoradiograph of Hinf I digested shad mtDNA separated on a 6% polyacrylamide gel. Variation can be seen in the presence/absence of 223 bp and 185 bp fragments. Lanes b, c, d, and f depict Genotype A, while lanes a, g, and h have Genotype B.

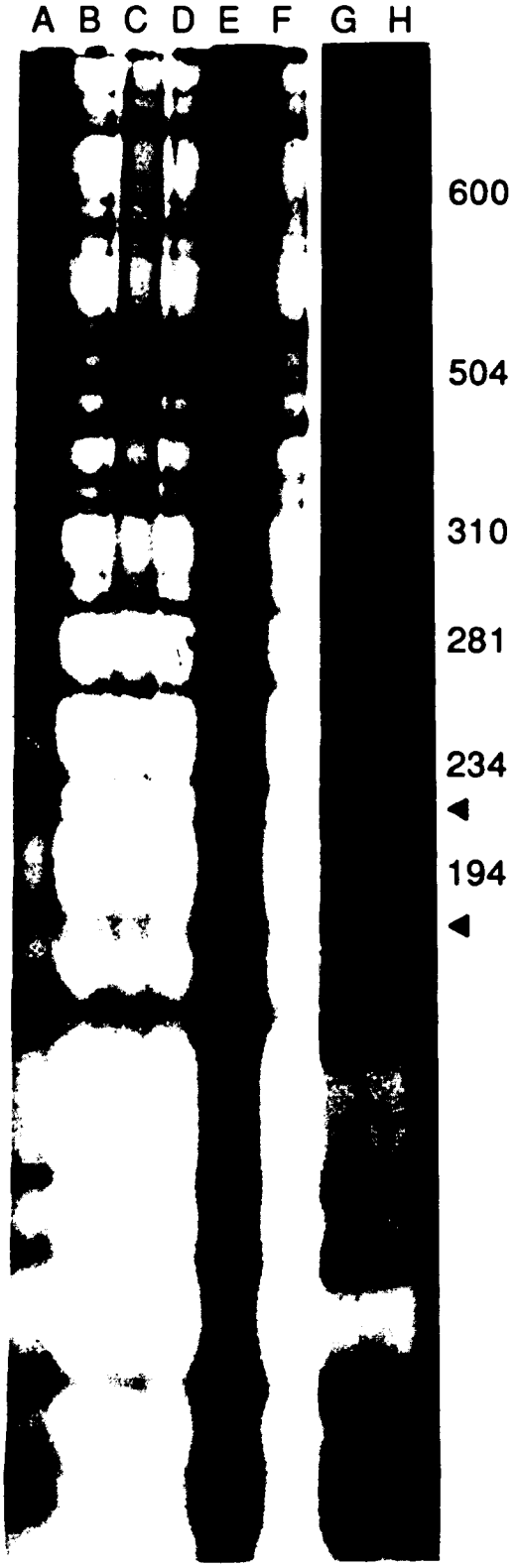

A B C D E F G H
600
504
310
281
234
194

Fig. 8. Representative Rsa I digests of shad mtDNA fractionated on a 6% polyacrylamide gel. Polymorphisms can be observed in the electrophoretic mobility of several mtDNA fragments which range from approximately 300 to 365 bp. Polymorphic region is indicated by the arrow. The molecular weight standard 1 kb ladder is in lane h. Genotype A is seen in lanes b, c, and d; Genotype B is in lane g; Genotype C is in lane a; and Genotype E is in lanes e and f.

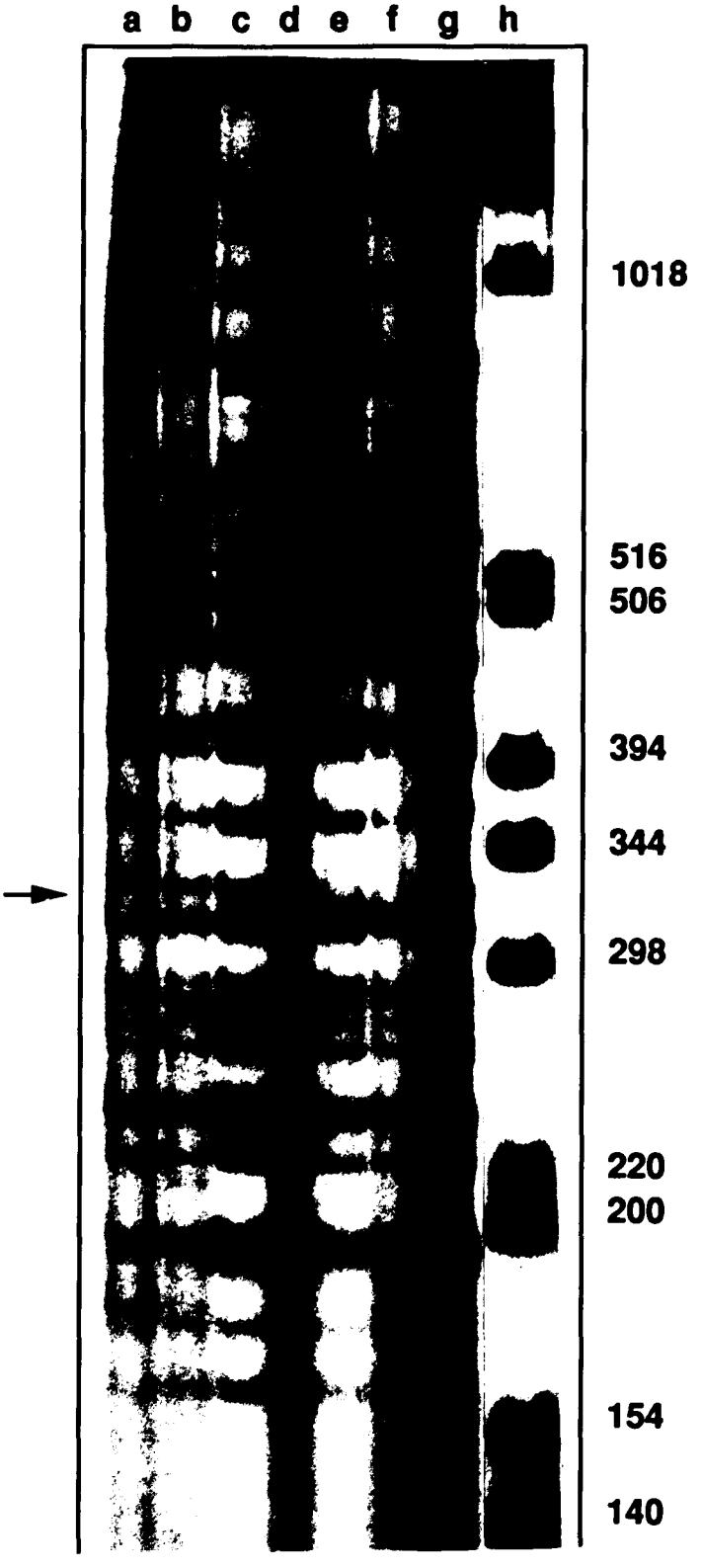
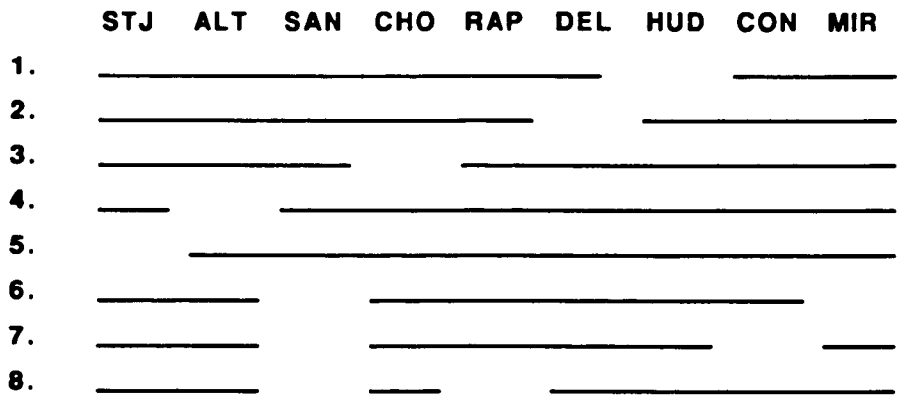

a b c d e f g h
1018
516
506
394
344
298
220
200
154
140

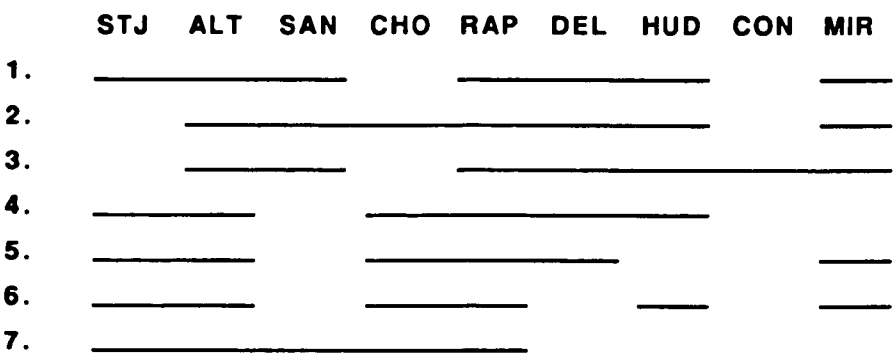
Fig. 9. Unplanned goodness-of-fit tests of homogeneity of individual restriction enzyme frequencies of shad mtDNA genotypes as revealed by (a.) Ava II, (b.) Dde I agarose and (c.) Rsa I gels. The bars indicate homogeneous subsets when rivers (spaces) are excluded from the analysis. Data presented are all genotypes scanned simultaneously by the BIOM R X C program (Sokal and Rohlf 1981).

Ava II



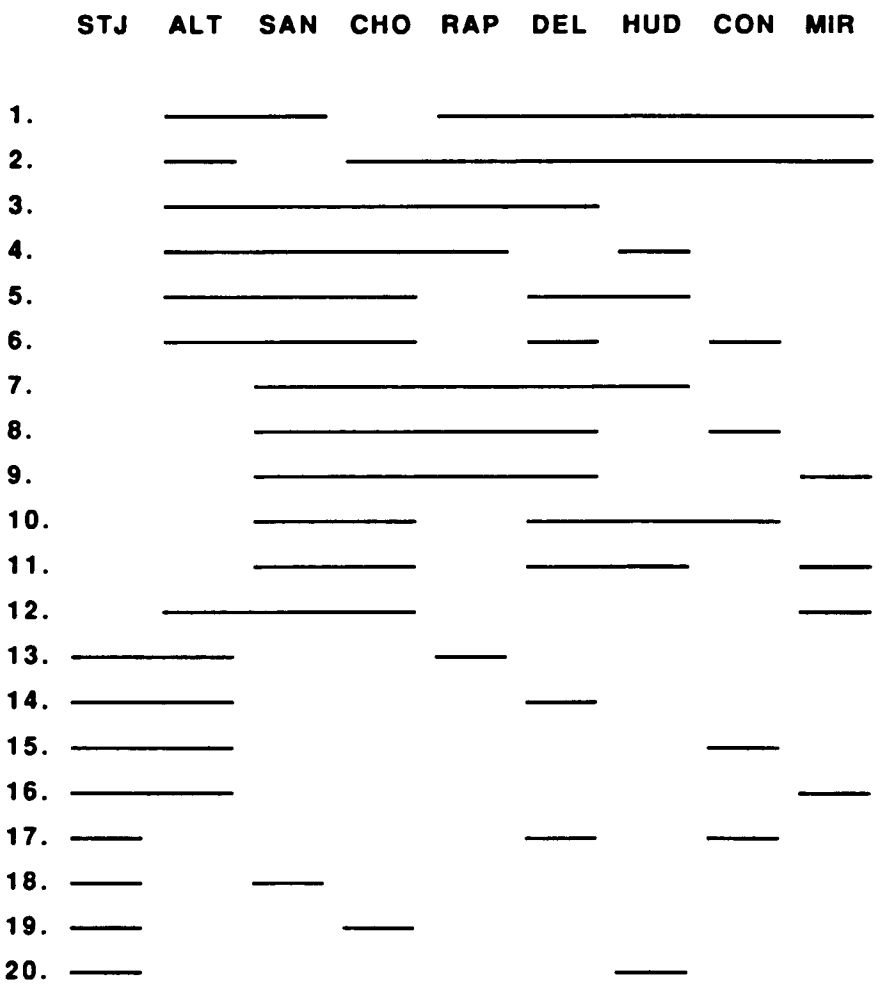
a.

Dde I



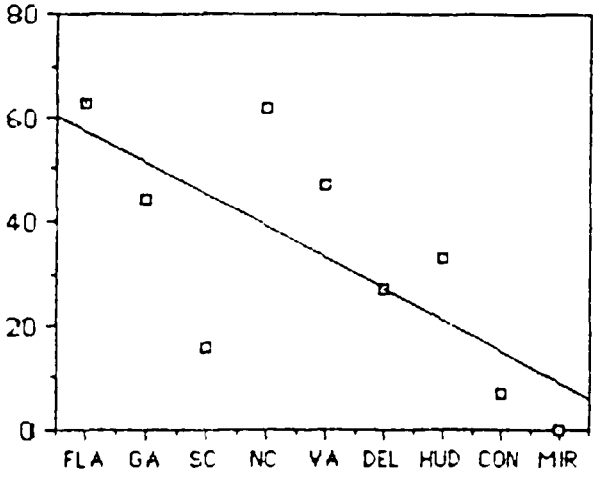
b.

Rsa I


STJ
ALT
SAN
CHO
RAP
DEL
HUD
CON
MIR
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14.
15.
16.
17.
18.
19.
20.

C.

Fig. 10. This figure depicts correlation of Dde I Genotype B with latitude. The x-axis represents the rivers and the y axis displays the frequencies as a percentage of the total number of fish sampled that river. The equation for the best straight line is given along with the significance of the R^2.



y = 63.556 - 6.0667x R^2 = 0.532

Fig. 11. Display of shared composite mtDNA genotypes among shad from different rivers. The bars indicate which rivers contain the shared composite mtDNA (numbered) genotypes in common.

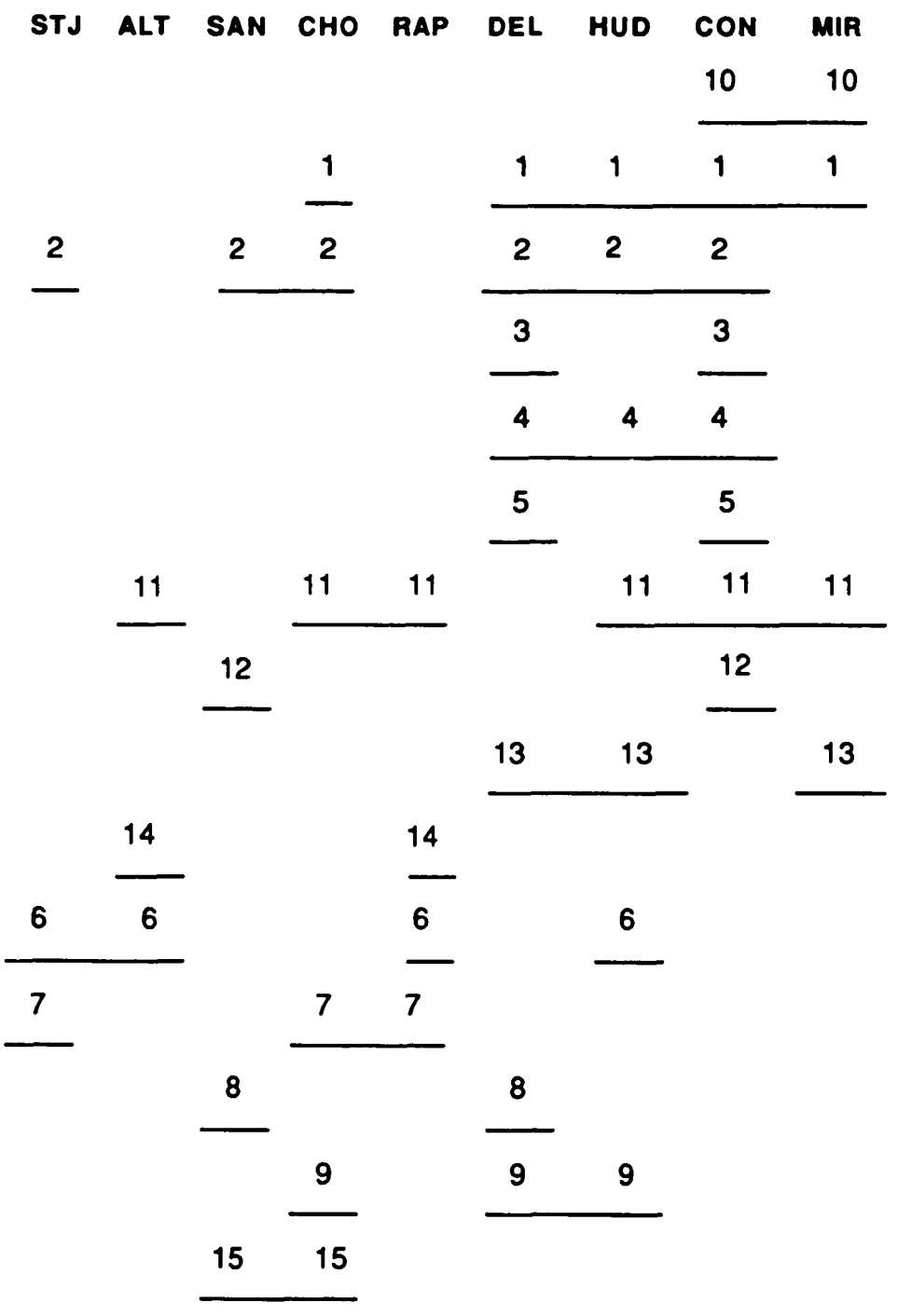
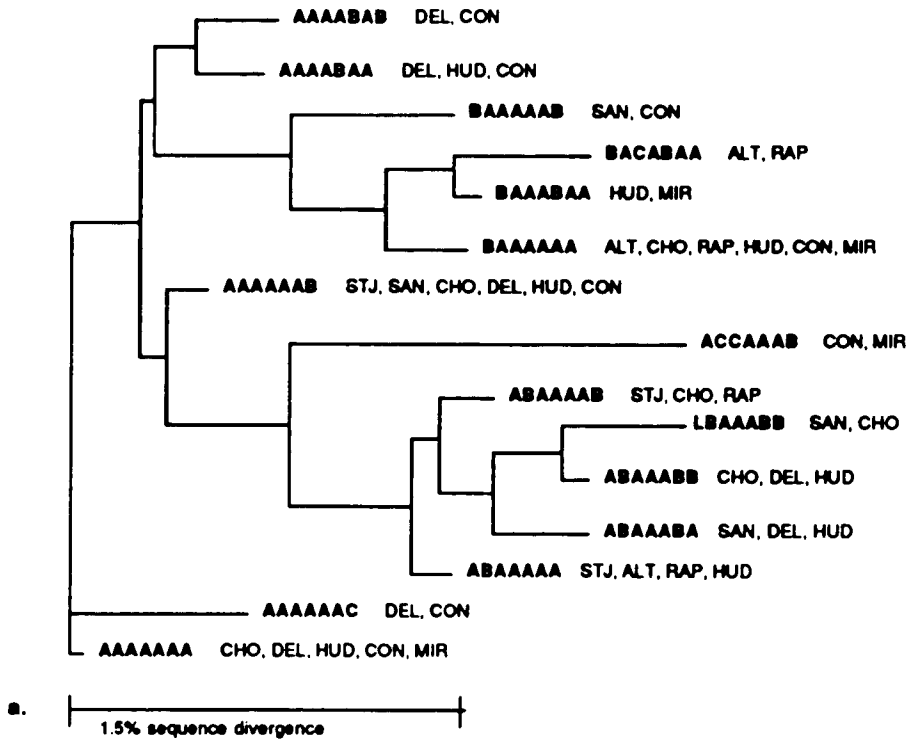

STJ
ALT
SAN
CHO
RAP
DEL
HUD
CON
MIR
10
10
1
1
1
1
2
2
2
2
2
2
3
3
4
4
4
5
5
11
11
11
11
11
11
12
12
13
13
13
14
14
6
6
6
6
7
7
7
8
8
9
9
9
15
15

Fig. 12. Fitch-Margoliash phenograms relating shared composite mtDNA genotypes. (among American shad from different Atlantic coastal rivers (a.) unrooted and (b.) rooted with hickory shad mtDNA genotypes.


AAAABAB DEL, CON
AAAABAA DEL, HUD, CON
BAAAAAB SAN, CON
BACABAA ALT, RAP
BAAABAA HUD, MIR
BAAAAAA ALT, CHO, RAP, HUD, CON, MIR
AAAAAAB STJ, SAN, CHO, DEL, HUD, CON
ACCAAAB CON, MIR
ABAAAAB STJ, CHO, RAP
LBAAABB SAN, CHO
ABAAABB CHO, DEL, HUD
ABAAABA SAN, DEL, HUD
ABAAAAA STJ, ALT, RAP, HUD
AAAAAAC DEL, CON
AAAAAAA CHO, DEL, HUD, CON, MIR
a.
1.5% sequence divergence

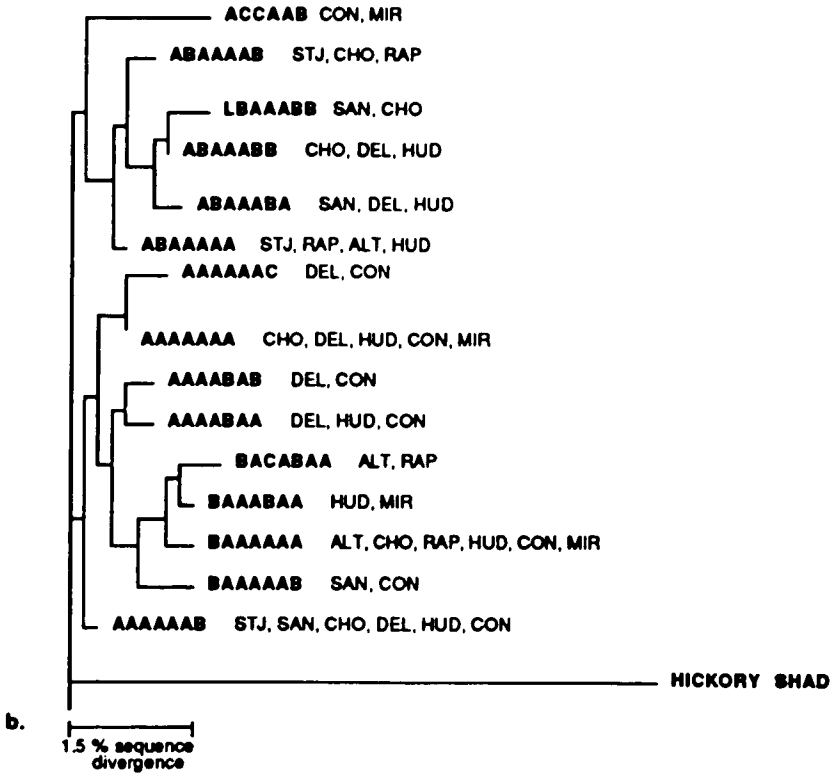
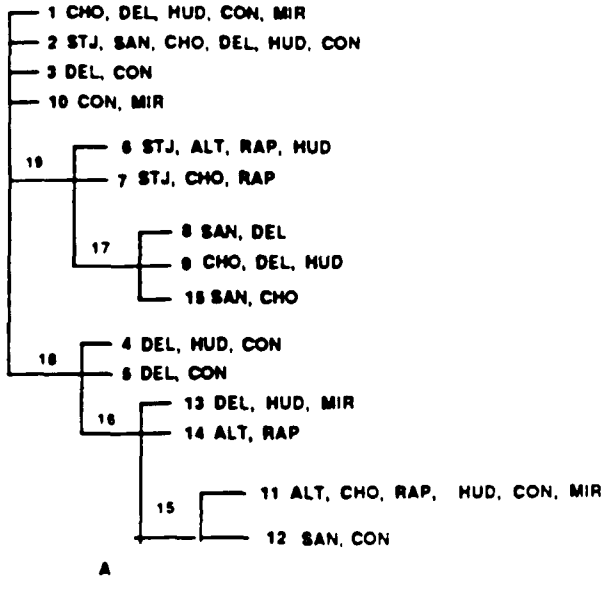
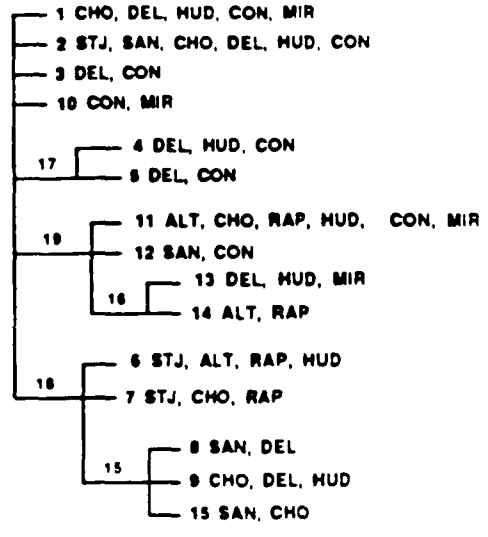

ACCAAB CON, MIR
ABAAAAB STJ, CHO, RAP
LBAAABB SAN, CHO
ABAAABB CHO, DEL, HUD
ABAAABA SAN, DEL, HUD
ABAAAAA STJ, RAP, ALT, HUD
AAAAAAC DEL, CON
AAAAAAA CHO, DEL, HUD, CON, MIR
AAAABAB DEL, CON
AAAABAA DEL, HUD, CON
BACABAA ALT, RAP
BAAABAA HUD, MIR
BAAAAAA ALT, CHO, RAP, HUD, CON, MIR
BAAAAAB SAN, CON
AAAAAAB STJ, SAN, CHO, DEL, HUD, CON
HICKORY SHAD
b.
1.5 % sequence divergence

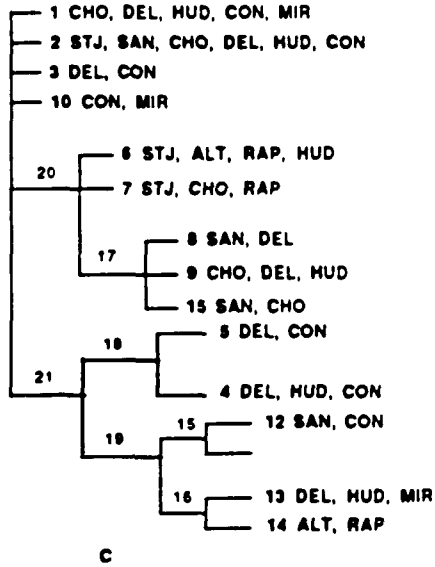
Fig. 13. The three most parsimonious unrooted trees for shared composite mtDNA genotypes using the HENNIG86 program (a, b, and c) and the most parsimonious tree for shared composite mtDNA genotypes rooted with hickory shad mtDNA genotypes (d). Numbers above branches indicate number of character state changes. Numbers at end of branches are the composite genotype number (out of 98). The CI of a, b, and c is 0.93 and of d is 0.96. The RI of all trees is 0.96.


1 CHO, DEL, HUD, CON, MIR
2 STJ, SAN, CHO, DEL, HUD, CON
3 DEL, CON
10 CON, MIR
19
6 STJ, ALT, RAP, HUD
7 STJ, CHO, RAP
17
8 SAN, DEL
9 CHO, DEL, HUD
15 SAN, CHO
18
4 DEL, HUD, CON
5 DEL, CON
16
13 DEL, HUD, MIR
14 ALT, RAP
15
11 ALT, CHO, RAP, HUD, CON, MIR
12 SAN, CON

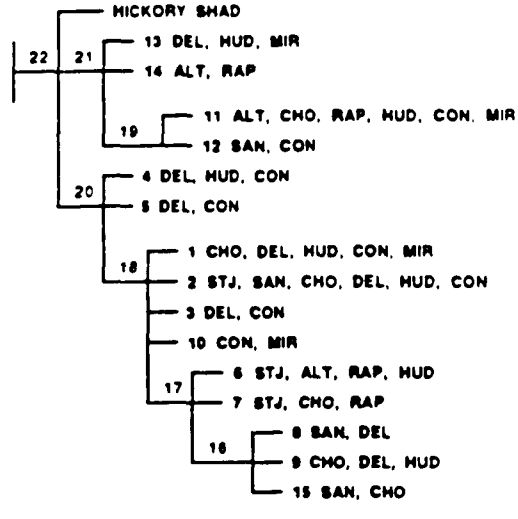
A


1 CHO, DEL, HUD, CON, MIR
2 STJ, SAN, CHO, DEL, HUD, CON
3 DEL, CON
10 CON, MIR
17
4 DEL, HUD, CON
5 DEL, CON
19
11 ALT, CHO, RAP, HUD, CON, MIR
12 SAN, CON
16
13 DEL, HUD, MIR
14 ALT, RAP
18
6 STJ, ALT, RAP, HUD
7 STJ, CHO, RAP
15
8 SAN, DEL
9 CHO, DEL, HUD
15 SAN, CHO

B


1 CHO, DEL, HUD, CON, MIR
2 STJ, SAN, CHO, DEL, HUD, CON
3 DEL, CON
10 CON, MIR
20
6 STJ, ALT, RAP, HUD
7 STJ, CHO, RAP
17
8 SAN, DEL
9 CHO, DEL, HUD
15 SAN, CHO
18
5 DEL, CON
21
4 DEL, HUD, CON
19
15
12 SAN, CON
16
13 DEL, HUD, MIR
14 ALT, RAP

C


HICKORY SHAD
22
21
13 DEL, HUD, MIR
14 ALT, RAP
19
11 ALT, CHO, RAP, HUD, CON, MIR
12 SAN, CON
20
4 DEL, HUD, CON
5 DEL, CON
18
1 CHO, DEL, HUD, CON, MIR
2 STJ, SAN, CHO, DEL, HUD, CON
3 DEL, CON
10 CON, MIR
17
6 STJ, ALT, RAP, HUD
7 STJ, CHO, RAP
16
8 SAN, DEL
9 CHO, DEL, HUD
15 SAN, CHO

D

LITERATURE CITED

Adams, J., and E. D. Rothman. 1982. Estimation of phylogenetic relationships from DNA restriction patterns and selection of endonuclease cleavage sites. Proc. Natl. Acad. Sci. USA 79:3560-3564.

Allendorf, F. W. 1983. Isolation, gene flow, and genetic differentiation among populations. pp. 51-65 in C. M. Schonewald-Cox, S. M. Chambers, B. MacBryde, and W. L. Thomas, eds. Genetics and Conservation. Benjamin Cummings Publishers, Menlo Park, CA.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. L. deBruijn, A. R. Coulson, J. Drouin, I. C. Eperon, D. P. Nierlich, B. A. Roe, F. Sanger, P. H. Schreir, A. J. H. Smith, R. Stadon, and I. G. Young. 1981. Sequence and organization of the human mitochondrial genome. Nature 290:457-465.

Araya, A., R. Amthauer, G. Leon, and M. Krauskopf. 1984. Cloning, physical mapping and genome organization of mitochondrial DNA from Cyprinus carpio oocytes. General Genetics 196: 43-52.

Avise, J. C. 1989. Gene trees and organismal histories: a phylogenetic approach to population biology. Evolution 43:1192-1208.

Avise, J. C. 1984. Characterization of mitochondrial DNA variability in a hybrid swarm between subspecies of bluegill sunfish (Lepomis macrochirus). Evolution 38:931-941.

Avise, J. C., J. Arnold, R. M. Ball, E. Bermingham, T. Lamb, J. E. Neigel, C. A. Reeb, and N. C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18:489-522.

Avise, J. C., B. W. Bowen, and T. Lamb. 1989. DNA fingerprints from hypervariable mitochondrial genotypes. Mol. Biol. Evol. 6:258-259.

Avise, J. C., G. S. Helfman, N. C. Saunders, and L. S. Hales. 1986. Mitochondrial DNA differentiation in North Atlantic eels: population genetic consequences of an unusual life history pattern. Proc. Natl. Acad. Sci. USA 83:4350-4354.

.Avise, J. C., and R. A. Lansman, 1983. Polymorphism of mitochondrial DNA in populations of higher animals. pp. 147-161. in M. Nei and R. K. Koehn eds. Evolution of Genes and Proteins. Sinauer Associates, Sunderland, MA.

Avise, J. C., C. A. Reeb, and N. C. Saunders. 1987. Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). Evolution 41:991-1002.

Avise, J. C., and N. C. Saunders. Hybridization and introgression among species of sunfish (Lepomis): analysis by mitochondrial DNA and allozyme markers. Genetics 108:237-255.

Baird, S. F. 1873-1889. Reports of the U. S. Fish Commission. Government Printing Office, Washington, D. C.

Barth, E. 1989. The State of the Chesapeake Bay. Third Annual Monitoring Report.

Berg, W. J., and S. D. Ferris. 1984. Restriction endonuclease analysis of salmonid mitochondrial DNA. Can. J. Fish. Aquat. Sci. 41:1041-1047.

Bentzen, P. 1989. Mitochondrial DNA polymorphism in American shad (Alosa sapidissima) and its implications for population structure. Ph. D. thesis, McGill University, Montreal, Quebec.

Bentzen, P., G. C. Brown, and W. C. Leggett. 1989. Mitochondrial DNA polymorphism, population structure, and life history variation in American shad (Alosa sapidissima). Can. J. Fish. Aquat. Sci. 46:1446-1454.

Bentzen, P., W. C. Leggett, and G. C. Brown. 1988. Length and restriction site heteroplasmy in the mitochondrial DNA of American shad (Alosa sapidissima). Genetics 118: 509-518.

Bermingham, E., and J. C. Avise. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. Genetics 113:939-966.

Bernatchez, L., J. J. Dodson, and S. Boivin. 1989. Population bottlenecks: influence on mitochondrial DNA diversity and its effect in coregonine stock discrimination. J. Fish Biol. 35:233-244.

Bernatchez, L., and J. J. Dodson. 1990. Mitochondrial DNA variation among anadromous populations of cisco (Coregonus artedii) as revealed by restriction analysis. Can. J. Fish. Aquat. Sci. 47:533-543.

Birt, T. P., J. M. Green, and W. S. Davidson. 1986. Analysis of mitochondrial DNA in allopatric, anadromous, and nonanadromous Atlantic salmon, Salmo salar. Can. J. Zoo. 64:118-120.

Blaxter, J. H. S. 1984. Ontogeny, systematics and fisheries. American Society of Ichthyologists and Herpetologists Special Publication 1:1-6.

Brandt, R. E. 1987. Current status of American shad, alewife, and blueback herring in New York waters. Prepared for the Atlantic States Marine Fisheries Commission shad and river herring scientific and statistical committee. Albany, NY.

Brown, W. M., M. George, and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76:1967-1971.

Buroker, N. E., J. R. Brown, T. A. Gilbert, P. J. O'Hara, A. T. Beckenbach, W. K. Thomas, and M. J. Smith. 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics 124:157-163.

Campbell, R. C. 1967. Statistics for Biologists. Cambridge University Press. Cambridge, UK.

Carscadden, J. E., and W. C. Leggett. 1975. Meristic differences in spawning populations of American shad, Alosa sapidissima: evidence for homing to tributaries in the St. John River, New Brunswick. J. Fish. Res. Bd. Can. 32:653-660.

Chapman, R. W. 1987. Changes in the population structure of male striped bass, Morone saxatilis, spawning in the three areas of the Chesapeake Bay from 1984 to 1986. U.S. Fish Wildl. Serv. Fish. Bull. 85:167-170.

Chapman, R. W. 1989. Spatial and temporal variation of mitochondrial DNA haplotypes frequencies in the striped bass (Morone saxatilis) 1982 year class. Copeia 1989:344-348.

Chapman, R. W. 1990. Mitochondrial DNA analysis of striped bass populations in Chesapeake Bay. Copeia 1990:355-366.

Clayton, D. A., J. N. Doda, and E. C. Friedberg. 1974. The absence of a pyrimidine dimer repair mechanism in mammalian mitochondria. Proc. Nat. Acad. Sci. USA 74:2777-2784.

Cooper, J. C., ed. 1984. Proceedings of a workshop on critical data needs for shad research on the Atlantic coast of North America. Hudson River Foundation for Science and Environmental Research, Inc., New York, NY.

Crecco, V. A., and T. F. Savoy. 1984. Effects of fluctuations in hydrographic conditions on year-class strengths of American shad (Alosa sapidissima) in the Connecticut River. Can. J. Fish. Aquat. Sci. 41:1216-1223.

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

Dadswell, M. J., R. A. Rulifson, and G. R. Daborn. 1986. Potential impact of large-scale tidal power developments in the upper Bay of Fundy on fisheries resources of the Northwest Atlantic. Fisheries 11:26-35.

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

Davidson, W. S., T. P. Birt, and M. Green. 1989. A review of genetic variation in Atlantic salmon, Salmo salar L., and its importance for stock identification, enhancement programmes and aquaculture. J. of Fish Biol. 35: 547-560.

DesFosse, J. C., and J. G. Loesch. 1985. Multivariate analysis of meristic characters of juvenile Alosa. Virginia Institute of Marine Sciences, Completion report AFC 15-1, Gloucester Point.

Densmore, L. D. , J. W. Wright, and W. M. Brown. 1985. Length variation and heteroplasmy are frequent in mitochondrial DNA from parthenogenetic and bisexual lizards (Genus Cnemidophorus). Genetics 110:689-707.

Dodson, J. J., and W. C. Leggett. 1973. Behavior of adult American shad (Alosa sapidissima) homing to the Connecticut River from Long Island Sound. J. Fish. Res. Board Can. 30:1847-1860.

Dowling, T. E., and W. M. Brown. 1989. Allozymes, mtDNA and levels of phylogenetic resolution among four minnow species (Notropis: Cyprinidae). Syst. Zool. 38:126-143.

Farris, J. S. 1969. A successive approximations approach to character weighting. Syst. Zool. 18:374-385.

Farris. J. S. 1989. The retention index and the rescaled consistency index. Cladistics 5:417-420.

Felsenstein, J. F. 1988. Phylogenies from molecular sequences: inference and reliability. Ann. Rev. Genet. 22:521-565.

Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. Science 155:279-284.

Fodor, R.V. 1981. Frozen Earth: Exploring the Ice Ages. Enslow Publ., Hillside, NJ.

Gibson, M. R., V. A. Crecco, and D. E. Stang. 1988. Stock assessment of American shad from selected Atlantic coast rivers. Special Report No. 15 of the Atlantic States Marine Fisheries Commission.

Glebe B. D., and W. C. Leggett. Temporal, intra-population differences in energy allocation and use by American shad (Alosa sapidissima) during the spawning migration. Can. J. Fish. Aquat. Sci. 38:795-805.

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

Grahl-Nielsen, O., and K. A. Ulvund. 1990. Distinguishing populations of herring by chemometry of fatty acids. American Fisheries Society Symposium 7:566-571.

Graves, J. E., and A. E. Dizon. 1989. Mitochondrial DNA sequence similarity of Atlantic and Pacific albacore tuna (Thunnus alalunga). Can. J. Fish. Aquat. Sci. 46:870-873.

Graves, J. E., S. D. Ferris, and A. E. Dizon. 1984. Close genetic similarity of Atlantic and Pacific skipjack tuna (Katsuwonus pelamis) demonstrated with restriction endonuclease analysis of mitochondrial DNA. Mar. Biol. 79:315-319.

Greeley, J. R. 1936. Fishes of the area with annotated list. p. 50 in A Biological Survey of the Lower Hudson Watershed. NY Conservation Dept. J. B. Lyons Co. Albany, NY.

Gyllensten, U., D. Wharton, and A. C. Wilson. 1985. Maternal inheritance of mitochondrial DNA during backcrossing of two species of mice. J. Hered. 76:321-324.

Gyllensten, U., and A. C. Wilson. 1987. Mitochondrial DNA of salmonids: inter and intraspecific variability detected with restriction enzymes. pp. 301-318. in N. Ryman and F. Utter. eds. Population Genetics and its Application to Fisheries Management. U. of Wash. Press. Seattle, WA.

Hildebrand, S. F. 1963. Family Clupeidae. Fishes of the Western North Atlantic. Mem. Sears Found. Mar. Res. 1(3): 257-454.

Hollis, E. H. 1948. The homing tendency of shad. Science 108:332-333.

Hynes, R. A., E. J. Duke, and P. Joyce. 1989. Mitochondrial DNA as a genetic marker for brown trout, Salmo trutta L., populations. J. Fish Biol. 35:687-701.

Johnson, J. R., J. G. Loesch, and Alan B. Blair. 1986. A morphometrical comparison between cultured and wild juvenile American shad. Progressive Fish Culturist 48:168-170.

Kluge, A. G., and J. S. Farris. 1969. Quantitative phyletics and the evolution of anurans. Syst. Zool. 18:1-32.

Kornfield, I., and M. Bogdanowicz. 1987. Differentiation of mitochondrial DNA in Atlantic herring, Clupea harengus. U. S. Fish Wildl. Ser. Fish. Bull. 85:561-568.

Lamb, T., J. C. Avise, and J. Whitfield-Gibbons. 1989. Phylogeographic patterns in mitochondrial DNA of the desert tortoise (Xerobates agassizi), and evolutionary relationships among the North American gopher tortoises. Evolution 43:76-87.

Lansman, R. A., J. C. Avise, and M. D. Huettel. 1983. Critical experimental test of the possibility of "paternal leakage" of mitochondrial DNA. Proc. Natl. Acad. Sci. USA 80:1969-1971.

Leaverton, P. E. 1986. A Review of Biostatistics. Little, Brown and Co. Boston.

Leggett, W. C. 1976. The American shad (Alosa sapidissima) with special references to its migration and population dynamics in the Connecticut River. pp. 169-225. in D. Merriman and L. M. Thorpe eds. The Connecticut River Ecological Study, the Impact of a Nuclear Power Plant. American Fisheries Society Monograph 1.

Leggett, W. C., and J. Carscadden. 1978. Latitudinal variation in reproductive characteristics of American shad (Alosa sapidissima): evidence for population specific life history strategies in fish. J. Fish. Res. Board Can. 35:1469-1478.

Leggett, W. C., and R. R. Whitney. 1972. Water temperatures and the migrations of American shad. U.S. Fish Wildl. Serv. Fish. Bull. 70:659-670.

Legrande, W. H., and J. M. Grady. 1991. Chromosome and allozyme variation among populations of the elegant madtom Noturus elegans (Siluriformes:Ictaluridae). abstract in American Society of Ichthyologists and Herpetologists. Program and Abstracts. American Museum of Natural History, NY, NY.

MacRae, A. F., and W. W. Anderson. 1988. Evidence for non-neutrality of mitochondrial DNA haplotypes in Drosophila pseudoobscura. Genetics 120:485-494.

Maniatis, T., E .F. Fritsch, and J. Sambrook. 1982. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor, NY.

Maniatis, T., E .F. Fritsch, and J. Sambrook. 1989. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor, NY.

Mansueti, R., and H. Kolb. 1953. A Historical Review of the Shad Fisheries of North America. Maryland Board of Natural Resources and Dept. of Research and Education. Chesapeake Biological Laboratory, Publ. No. 97. Solomons, MD.

Melvin, G. D. 1984. The usefulness of meristic and morphometric characters in discriminating populations of American shad (Alosa sapidissima) inhabiting Cumberland Basin, New Brunswick. Doctoral dissertation. University of New Brunswick, Fredericton, Canada.

Melvin, G. D., M. J. Dadswell, and J. D. Martin. 1986. Fidelity of American shad, Alosa sapidissima, to its river of previous spawning. Can. J. Fish. Aquat. Sci. 43:640-646.

Nei, M., and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76:5269-5273.

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

Neves, R. J., and L. Depres. 1979. The oceanic migration of American shad, Alosa sapidissima, along the Atlantic coast. U. S. Fish Wildl. Ser. Fish. Bull. 77:199-212.

Palva, T. K. 1986. Cytogenetic and mitochondrial DNA analysis of four salmonid fish species. Ph. D. thesis, Dept. of Applied Zoology and Dept. of Physiology, University of Kuopio, Kuopio, Finland.

Patterson, C. 1989. Phylogenetic relations of major groups: conclusions and prospects. pp. 471-478. in Fernholm, B., K. Brenner, and H. Jornvall eds. The Hierarchy of Life. Elsevier. New York.

Richkus, W.A., and G. DiNardo. 1984. Current status and biological characteristics of the anadromous alosid stocks of eastern United States: American shad, hickory shad, alewife, and blueback herring. Phase I in interstate management planning for migratory alosids of the Atlantic coast. Martin Marietta, Baltimore, MD. 262 pp.

Riddle, B. R., and R. L. Honeycutt. 1990. Historical biogeography in North American arid regions: an approach using mitochondrial DNA phylogeny in grasshopper mice (Genus Onychomys). Evolution 44:1-15.

Savoy, T. F., and V. A. Crecco. 1988. The timing and significance of density-dependent and density-independent mortality of American shad, Alosa sapidissima. U. S. Fish Wildl. Ser. Fish. Bull. 86:467-482.

Scott, W. B., and W. J. Crossman. 1973. Freshwater fishes of Canada. Bull. Fish. Res. Board Can. 184:966 pp.

Sidell, B. D., R. G. Otto, D. A. Powers, M. Karweit, and J. Smith. 1980. Apparent genetic homogeneity of spawning striped bass in the Upper Chesapeake Bay. Trans. Am. Fish. Soc. 109:99-107.

Singh, G., N. Neckelman, and D. C. Wallace. 1987. Conformational mutations in human mitochondrial DNA. Nature 329:270-273.

Sismour, E.N., and R.S. Birdsong. 1986. A biochemical and genetic analysis of American shad (Alosa sapidissima) migrating into the Chesapeake Bay. American Fisheries Society Symposium 1:555.

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

Slatkin, M. 1989. Detecting small amounts of gene flow from phylogenies of alleles. Genetics 121:609-612.

Slatkin, M., and W. P. Maddison. 1989. A cladistic measure of gene flow. Genetics 123:603-613.

Smith. P. J., A. J. Birley, A. Jamieson, and C. A. Bishop. 1989. Mitochondrial DNA in the Atlantic cod, Gadus morhua: lack of genetic divergence between eastern and western populations. J. of Fish Biol. 36:369-373.

Sokal, R. R. and F. J. Rohlf. 1969. Biometry. W. H. Freeman, San Francisco.

Sokal, R. R. and F. J. Rohlf. 1981. Biometry. W. H. Freeman, San Francisco.

Stevens, D. E., H. K. Chadwick, and R. E. Painter. 1987. American shad and striped bass in California's Sacramento-San Joaquin River system. American Fisheries Society Symposium 1:66-78.

Summers, J. K. and K. A. Rose. 1987. The role of interactions among environmental conditions in controlling historical fisheries variability. Estuaries 3:255-266.

Talbot, G. B., and J. E. Sykes. 1958. Atlantic coast migrations of American shad. U. S. Fish Wildl. Ser. Fish. Bull. 58:473-490.

Templeton, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221-244.

Thomas, W. K., R. E. Withler, and A. T. Beckenbach. 1986. Mitochondrial DNA analysis of Pacific salmonid evolution. Can. J. Zool. 64:1058-1064.

Upholt, W. B. 1977. Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res. 4:1257-1265.

Wagner, G. P. 1989. The biological homology concept. Annu. Rev. Ecol. Syst. 20:51-69.

Walburg, C. H. and R. P. Nichols. 1967. Biology and management of the American shad and status of the fisheries, Atlantic coast of the United States, 1960. U.S. Fish Wildl. Serv. Spec. Sci. Rep. Fish 550:1-105.

Waldman, J. 1991. Disagreement among stock identification techniques for striped bass: blame it on the phenotype? abstract in American Society of Ichthyologists and Herpetologists. Program and abstracts. American Museum of Natural History, New York, NY.

Wardlaw, A. C. 1985. Practical Statistics for Experimental Biologists. Wiley, New York. 290 pp.

Whitehead, P. J. 1984. Fishes of the north-eastern Atlantic and the Mediterranean. UNESCO, Paris. 268-281.

Williams, P. J. 1985. Use of otoliths for stock differentiation of American shad (Alosa sapidissima Wilson). Masters thesis. Acadia University, Wolfville, Canada.

Williams, R. O., and G. E. Bruger. 1972. Investigations on American shad in the St. Johns River. Technical Series No. 66. Florida Dept. of Natural Resources, Marine Research Laboratory, St. Petersburg, FLA.

Williams, R. G., and G. Daborn. 1984. Spawning of the American shad in the Annapolis river, Nova Scotia, Canada. Proc. N.S. Institute Sci. 34:9-14.

Wirgin, I. I. 1987. Molecular evolution in the fish genus Morone. Ph. D. Thesis. City University of New York, NY.

Wirgin, I. I., and L. Maceda. 1991. A comparison of nuclear and mtDNA polymorphisms among striped bass populations. abstract in American Society of Ichthyologists and Herpetologists. Program and abstracts. American Museum of Natural History, New York, NY.

Wirgin, I. I., R. Proenca, and J. Grossfield. 1989. Mitochondrial DNA diversity among populations of striped bass in the southeastern United States. Can. J. of Zoo. 67:891-907.

Wirgin, I. I. , M. D'Amore, C. Grunwald, A. Goldman, and S. J. Garte. 1990a. Genetic diversity at an oncogene locus and in mitochondrial DNA between populations of cancer-prone Atlantic tomcod. Biochem. Genet. 28:459-475.

Wirgin, I. I. P. Silverstein, and J. Grossfield. 1990b. Restriction endonuclease analysis of striped bass mitochondrial DNA: the Atlantic coastal migratory stock. American Fisheries Symposium 7:475-491.

Wodtke, E. 1974. Eigenschaften von isolierten mitochondrien des gales und ihre abhagigkeit von der adaptations temperatur unter besonderer berucksichtigung der oxiativen phosphorylierung. J. Comp. Physiol. 91: 277-307.

Zink, R. M. and J. C. Avise. 1990. Patterns of mitochondrial DNA and allozyme evolution in the avian genus Ammodramus. Syst. Zool. 39:148-161.