FINE SCALE POPULATION STRUCTURE OF BAY SCALLOPS IN NORTH CAROLINA REVEALED BY AFLP FINGERPRINTING

Mark W. Sherman

A Thesis Submitted to the University of North Carolina Wilmington in Partial Fulfillment of the Requirements for the Degree of Master of Science

Department of Biology and Marine Biology
University of North Carolina Wilmington
2010

Approved by

Advisory Committee

Marcel van Tuinen Michael A. McCartney

Ami E. Wilbur Chair

Accepted by

Dean, Graduate School
TABLE OF CONTENTS

ABSTRACT ........................................................................................................ iv
ACKNOWLEDGMENTS ...................................................................................... v
LIST OF TABLES ............................................................................................ vi
LIST OF FIGURES ........................................................................................... vii
INTRODUCTION ............................................................................................... 1
  Overview ........................................................................................................ 1
  Objective ........................................................................................................ 11
  Significance .................................................................................................. 12
METHODS ......................................................................................................... 13
  Sample Collection ...................................................................................... 13
  DNA Extraction .......................................................................................... 14
  AFLP Analysis ............................................................................................ 15
  Fragment Detection and Signature Evaluation ........................................... 16
  Statistical Analysis ....................................................................................... 18
RESULTS ........................................................................................................... 21
  Descriptive Statistics ................................................................................ 21
  Population Comparisons between New York and North Carolina ........... 24
    Population Comparisons within North Carolina ...................................... 28
DISCUSSION .................................................................................................... 34
CONCLUSIONS ................................................................................................ 39
LITERATURE CITED ......................................................................................... 40
ABSTRACT

In recent years, bay scallops (*Argopecten irradians*) in North Carolina have been the cause of growing concern due to decreased abundance. The decline is attributed to a combination of natural and anthropogenic factors that have upset natural connections between local aggregations once sufficient to keep bay scallops abundant throughout North Carolina. This study utilized amplified fragment length polymorphism analysis (AFLP), a whole genome analytical technique, to characterize the amount of genetic differentiation in scallops from eight sites in North Carolina (NC) and one in New York (NY). Six primer combinations yielded a total of 469 fragments. All 302 scallops analyzed exhibited unique fragment profiles. Analysis of molecular variance revealed significant differentiation between NC and NY ($\Phi_{ST}=0.0625$, $P<0.0001$) and among NC populations. Assignment tests correctly assigned 33 of 35 NY scallops back to NY and all 267 NC scallops back to NC. Pairwise comparisons among NC populations yielded $\Phi_{ST}$ values ranging from 0.0011 to 0.0246 (20/28 comparisons significant, $\alpha=0.05$). Within NC, a moderate proportion of scallops were correctly assigned to source populations; however accuracy increased markedly when collection sites were combined based on geographic proximity. Overall, the results suggest subtle, but significant, genetic structure in North Carolina bay scallop populations.
ACKNOWLEDGMENTS

Thanks to Dana Schmidt, Don Murdoch, Anne Markwith, Troy Alphin, Cassie Martin, and “the guy I met at the breakfast place in Ocracoke at 4:30am who told me where the scallops were” for field assistance.

I’d like to thank Victor Schmidt for struggling through exploration of the AFLP technique with me and technical assistance in the lab. I’d also like to thank Jenn Bagwell, Mike Escheverria, Jan Vicente, and Julie Campbell for companionship during analysis runs.

I’d like to thank my committee, Dr. Marcel van Tuinen for advice and encouragement during experimental design and analysis and Dr. Michael McCartney for introducing me to theoretical population genetics and facilitating the method and interpretation of the results. I thank my advisor, Dr. Ami E. Wilbur for providing me with encouragement, making time available, permitting me to share my results at conferences, allowing me to work with other techniques throughout my time at UNCW and for offering hope when methods, ideas or results did not seem sensible.

I’d like to thank my friends, family, and Caitlin McKinstry for encouragement and support.
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Scallops collected for genetic analysis and site abbreviations</td>
<td>13</td>
</tr>
<tr>
<td>2. AFLP fragment summary statistics by primer pair</td>
<td>21</td>
</tr>
<tr>
<td>3. Arlequin AMOVA results with groupings of {HC, NR, HB, WB, EB, HI, BI, OI} {NY}</td>
<td>27</td>
</tr>
<tr>
<td>4. Pairwise $\Phi_{ST}$ and significance matrix from Arlequin including New York after 10100 permutations</td>
<td>27</td>
</tr>
<tr>
<td>5. STRUCTURE output of 13 replicate North Carolina runs at $K=4$</td>
<td>30</td>
</tr>
<tr>
<td>6. AMOVA results of North Carolina populations grouped by geographic regions</td>
<td>31</td>
</tr>
<tr>
<td>7. AFLPOP reassignment test (leave-one-out) distribution of North Carolina populations</td>
<td>33</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sampling sites</td>
<td>13</td>
</tr>
<tr>
<td>2. The average number of fragments per site for each primer pair</td>
<td>22</td>
</tr>
<tr>
<td>3. Frequency distributions of the average number of fragments for each primer pair</td>
<td>23</td>
</tr>
<tr>
<td>4. Frequency distribution of total fragments across all primer pairs</td>
<td>23</td>
</tr>
<tr>
<td>5. STRUCTURE output of triplicate runs from K=1-6 including New York</td>
<td>25</td>
</tr>
<tr>
<td>6. Graphical description of implementation of the Evanno method including New York</td>
<td>25</td>
</tr>
<tr>
<td>7. Q-statistics including New York at K=5</td>
<td>26</td>
</tr>
<tr>
<td>8. STRUCTURE output of North Carolina triplicate runs from K=1-6</td>
<td>29</td>
</tr>
<tr>
<td>10. Q-statistics of North Carolina populations at K=4</td>
<td>30</td>
</tr>
<tr>
<td>11. Regional allocation of assignment tests within North Carolina</td>
<td>33</td>
</tr>
</tbody>
</table>
INTRODUCTION

Overview

The bay scallop, *Argopecten irradians*, is a suspension feeding, epibenthic bivalve found in seagrass beds ranging from the Cape Cod area at its northern limit to central Mexico at its southern limit (Marelli *et al.* 1997). Historically, bay scallops have provided both social and economic importance to coastal towns throughout the Eastern United States. These bivalves were plentiful in coastal areas and utilized in local food, jewelry, and art. After the civil war, demand for scallops grew in non-coastal areas. Scallops that were once only consumed locally were transported inland to restaurants and seafood distributors (Gutsell 1928). With increased transport of bay scallops came heightened demand. In the years following the distribution of bay scallops, and even to date, the scallop industry has not been able to satisfy the demands of consumers. The lucrative fishery rapidly depleted bay scallop stocks along the coast and population levels began fluctuating wildly on a yearly basis (Goldberg & Tettelbach 2008, Leverone *et al.* 2008, Chestnut 1951). In recent years, environmental factors such as hurricanes, habitat degradation, increased predation, algal blooms, and loss of submerged aquatic vegetation (SAV) habitat have also likely contributed to the loss of scallop stock (Gutsell 1930, Tettelbach *et al.* 1985, Peterson *et al.* 1989, Summerson & Peterson 1990).

When marine species have reached threatened or critical population levels as bay scallops have, several strategies can be implemented to allow recovery and increase abundance. The least intrusive of these strategies, from the perspective of the animal, is to prohibit fishing and implement protection for the species. The imposition of a moratorium on the fishery releases the species from fishing pressures which could be a significant contributor to the suppression of the population. Another minimally intrusive restoration strategy is habitat restoration, a process by
which a limited substrate is provisioned and/or the impacts of coastal development on water quality are reduced. Substrate supplementation is most effective for organisms that are heavily reliant on a limited substrate, such as oysters (Mann et al. 1990). Submerged aquatic vegetation (SAV) has been supplemented to increase available nursery habitat for many important fishery species (Lee & Park 2008). Both strategies (fishing prohibition and habitat restoration) are aimed at reducing human impacts and rely on natural productivity (reproduction) to rebuild populations.

If low-intrusion strategies are unable to restore a stock, other strategies that directly impact the species of interest can be considered. Supplementation is a form of restoration that involves the augmentation of natural populations and/or natural processes to boost population numbers. The intention of supplementation is to raise the population abundance beyond a critical number of individuals at which the population becomes self-sustaining. Transplantation and hatchery-based augmentation are two types of supplementation currently used in restoration or enhancement efforts and differ primarily in the origin of the animals used to increase target population abundance. Transplantation involves the relocation of organisms from an area of high abundance to an area of low abundance and thus relies on natural productivity. Hatchery-based augmentation involves the production of animals in a controlled setting and their release into targeted restoration areas.

In the Eastern United States, several states have addressed the dwindling bay scallop stock with restoration attempts with varied success. Northeastern states have generally utilized supplementation strategies, relying on commercial aquaculture to maintain a healthy bay scallop industry. These approaches have resulted in successful recovery from recurring brown tide events in the Peconic Bays of New York (Tettelbach & Wenczel 1993, Goldberg & Tettelbach
In Florida, bay scallop population levels had decreased so drastically by the 1970s that local populations in areas such as Tampa Bay were extirpated from the region. The loss triggered the inception of a restoration program that has since been in existence for over thirty years. The program began with efforts to maintain the natural connections between populations in the Tampa Bay area. Bay scallops, raised by waterfront residents, were moved into Tampa Bay to fill the gap between the adjacent populations. The program further developed to include hatchery-based supplementation in which local scallops were spawned and their larvae reared until they were juveniles. The juveniles were deployed at high-density in cages to increase the chance of successful fertilization and enhance larval production when they reached adulthood (Arnold et al. 2005, Leverone et al. 2008). The strategy showed great short term success. Efforts made by this strategy increased local abundances by one order of magnitude the first year and another order of magnitude the next (Arnold & Blake 2006). After the second year of spawning and completion of the experimental supplementation, scallops declined back to background levels (Arnold & Blake 2006). In recent years, the program has shifted restoration strategies to include supplementation by hatchery production and the release of scallops at the pediveliger larval stage (Leverone et al. 2008). The efforts have resulted in higher abundance of scallops in Pine Island Sound, Tampa Bay, and St. Andrews Bay than before restoration efforts were taken (Leverone et al. 2008).

In North Carolina, scallop abundance was historically variable, but self-sustaining, until a six-week red tide event in 1987. The event caused a crash of the bay scallop population in North Carolina to about 10 to 15 percent of the historic (25 year average) abundance, from which it did not fully recover (Peterson & Summerson 1992). The crash of the fishery caused the North Carolina Department of Marine Fisheries (NCDMF) to implement a fishing moratorium on bay
scallops in 2001 and prompted the development of a bay scallop Fisheries Management Plan (FMP). Since its closure, NCDMF has regulated the bay scallop fishery and has performed yearly monitoring surveys to track population levels. The closure of the fishery has not allowed bay scallop populations to rebuild to historic levels, suggesting that natural production is inadequate to rebuild the populations or non-fishery related losses are sufficient to keep abundances low (NCDMF 2007). The potential for habitat degradation as a contributor to the continued low abundance has been considered. Several studies have highlighted the importance of SAV beds on scallop recruitment and development (Irlandi et al. 1995, Smith et al. 1989). For this reason, the protection and restoration of seagrass beds remain a high priority for NCDMF and have been addressed in the NC Department of Environmental and Natural Resources’ Coastal Habitat Protection Plan. However, these efforts to improve habitat have also failed to contribute to the substantial recovery of bay scallops, leading to the consideration of more invasive supplementation strategies to enhance scallop abundance.

To date, two experiments have investigated the effectiveness of supplementation strategies of bay scallops in North Carolina. One experiment involved transplantation of scallops from Oscar Shoal to Back Sound and Western Bogue Sound. The impact of the transplantations was quantified by the estimated change in local scallop recruit density, both to the local seagrass bed and to adjacent spat traps over subsequent years (Peterson & Summerson 1992, Peterson et al. 1996). The experiment was unable to find a significant increase of bay scallop recruitment at either site until scallops were supplemented to seagrass beds in high densities over several seasons. The investigators observed a 568% increase in scallop recruit density, suggested to be due to their supplementation efforts, at the receiver sites relative to before transplantation. The control sites, where no transplantation had occurred, only saw a 34% increase in scallop
abundance over the same time frame. The number of spat that recruited to the spat traps adjacent to the target site was expected to be commensurate with the number of spat recruited to the seagrass, but the study did not find such a correlation. The investigators did find that scallop recruitment to spat bags decreased as the distance from the release site increased (Peterson et al. 1996). This finding led to the conclusion that recruitment limitation is the reason that the North Carolina scallop population has failed to rebound to historic abundances. The other supplementation experiment involved hatchery-production of scallops and subsequent release of pediveliger larvae onto healthy seagrass beds in Western Bogue Sound, NC. The impacts were assessed by evaluating scallops collected 9 months after introduction for genetic signatures indicative of the hatchery spawns (AE Wilbur, UNCW, personal communication). While the mtDNA data suggested a modest augmentation (increased frequencies of haplotypes exhibited by the hatchery broodstock) the contribution was not supported by the analysis of microsatellites (genotypes inconsistent with scallops having a hatchery origin).

The increasing use of supplementation in the management of North Carolina’s bay scallop resource highlights the need for an understanding of the genetic structure of this species. Both transplantation and hatchery-based augmentation have the potential to alter natural patterns of genetic diversity and thus it is imperative to evaluate that diversity before any large scale supplementations are executed. Failure to do so may lead to homogenization (through transplantation) of disparate populations that are adapted to their native environments, restoration efforts targeted toward healthy populations, or hatchery-reared introduction of hundreds to thousands of offspring of very similar genetic composition into a naturally genetically diverse population. In time, the genes of the altered population may not be able to withstand environmental stresses that the organisms in the original population were adapted to tolerate,
rendering the restoration effort ineffective. Thus, a firm understanding of patterns of genetic diversity among bay scallop populations in North Carolina is imperative.

An understanding of bay scallop population dynamics in North Carolina is also important in the management of this resource. Populations of bay scallops are dispersed among a series of semi-enclosed bays and sounds from the lower Pamlico to the Cape Fear River. Because seagrass is patchily distributed throughout coastal environments, populations are discontinuously distributed within their habitat. If this patchiness limits larval exchange between populations, differentiation could occur on local levels. The possibility of differentiation due to isolation of populations is enhanced by the limited mobility of adults, relatively short pelagic larvae stage, and short lifespan (1-2 years) (Fay et al. 1983). If the disjunct nature of this distribution reflects isolation among aggregations, and if these populations are largely self-recruiting, then management plans need to focus on each discrete population. However, if there is high connectivity between sounds and scallops function as one interbreeding population, management can be employed to focus on the bay scallop population as a single unit. In addition, if particular areas are supplying other regions with larvae via water currents, it may be beneficial to cease fishing in these “source” areas to maximize the number of larvae produced and focus harvest effort in those “sink” areas that contribute little to subsequent generations (Peterson & Summerson 1992). One study found evidence of a “sink” population at Oscar Shoals in Back Sound, NC by showing that the exclusion of cow nose rays to an area of high predation allowed significantly increased survival of scallops in the enclosure (Peterson et al. 2001). By elucidating the population structure of bay scallops, an effective management plan or restoration strategy for North Carolina’s bay scallops can be developed.
The use of genetic techniques can provide inexpensive, powerful, and relatively rapid methods to identify genetically differentiated stocks. Several studies have attempted to determine the level of connectivity between geographic populations of broadcast spawners. If individuals in one geographic area interbreed with those in another geographic area, they may be deemed a single genetic population because they act as one interbreeding population. Since broadcast-spawning, sessile marine species rely on larval dispersal and currents to provide this connectivity, the influence that currents impose on larval dispersal warrants investigation. Significant genetic differentiation has been detected in studies involving sea urchins, and sea scallops, and coral (Banks et al. 2007, Dai et al. 2000, Kenchington et al. 2006). Banks et al. (2007) found significant fine scale genetic differentiation in Western Australia populations of the sea urchin, Centrostephanus rodgersii, amidst broad scale panmixia at much larger spatial scales. This pattern was attributed to the urchin’s long-distance dispersal in combination with local water currents isolating more nearby populations. Kenchington et al. (2006) found significant genetic differentiation in populations of the sea scallop, Placopecten magellanicus, in the Northwest Atlantic based on six microsatellite markers. The authors of this study concluded that the pattern of differentiation observed was likely due to tidal and current gradients and gyres associated with land that could retain planktonic larvae in relatively small areas. Dai et al. (2000) conducted an allozyme analysis of eight enzyme systems and found significant differentiation between populations of the scleractinian coral, Mycedium elephantotus, in North and South regions of Taiwan. Within the southern region, two genetically distinct populations were observed to be living in sympatry and it was suggested that the differentiation was due to asynchronous timing of gametogenesis and spawning. These studies suggest that the causes of
genetic differentiation can be attributed to a variety of mechanisms, working individually or in combination with one another.

In North Carolina, bay scallop population structure has been investigated by microsatellite and mitochondrial DNA analytical techniques (Hemond 2006, Marko & Barr 2007). These studies have shed some light on the genetic structure of the area, however, the two approaches led to somewhat contradictory conclusions. Hemond (2006) investigated nine microsatellite markers to determine the level of connectivity between scallop populations in Bogue and Core Sounds and observed no significant differentiation from the null hypothesis of homogeneity. A study utilizing mtDNA markers, however, observed significant genetic differentiation among scallops collected in the Intracoastal Waterway adjacent to Topsail Island, Bogue, Back, and Pamlico Sounds (Marko & Barr 2007). Conclusions based on this study suggest that each of these sounds contain genetically differentiated populations. The results of these studies propose two slightly opposing perspectives of bay scallop structure in North Carolina which may be partially attributable to the particular genetic markers used in these analyses.

Analysis of mtDNA is a powerful tool for uncovering population differentiation and is a common tool used in the delineation of fishery stocks (Griffiths et al. 2009, Yoon et al. 2008). Many studies have uncovered substantial differentiation between geographically proximal populations (Saunders et al. 1986, Liu et al. 2009). Saunders et al. (1986) found significant genetic differentiation between horseshoe crab populations across Cape Canaveral, FL utilizing mitochondrial restriction fragment length polymorphism analysis. The differentiation was suggested to be due to a barrier created by the offshore flow of water driven by the Gulf Stream. Liu et al. (2009) found deep significant differentiation between Northern and Southern regions of
China within grey mullets, *Mugil cephalus*. Within the large-scale differentiation, smaller spatial scales showed shallow, yet significant differentiation between populations.

While often useful for delineation of static populations, mtDNA markers are maternally inherited and do not respond as nuclear markers do when populations are isolated from one another. Thus, if two populations are isolated, the frequency of mitochondrial DNA and nuclear DNA may differentiate at different rates and cause discordance from other markers due to drift, mutation, migration, biased sex ratios, secondary contact, or mechanisms of reproductive isolation (Egger *et al.* 2007, Gonzalez & Zardoya 2007, Lemaire *et al.* 2005, Peijnenburg *et al.* 2006). Therefore, the use of alternative genetic markers in addition to mtDNA can be useful for resolving true population structure.

The utilization of microsatellites in population differentiation analyses is thought to provide a more comprehensive assessment than mtDNA because the technique allows the comparison of multiple independent, co-dominant nuclear loci (Hemond 2006) allowing for greater statistical power in detecting differentiation than is possible using mtDNA analysis. However, like most genetic techniques, microsatellites are not without limitations. Microsatellite analyses of closely related populations rely on a balance between genetic drift and mutation to differentiate populations over time. Microsatellites arise via slippage mutations through meiotic recombination or DNA misrepair, leading to insertions or deletions of tandem repeats (Tautz & Schlötterer 1994). If these mutations arise independently of one another, the same number of repeats may not be identical by descent, but will be analyzed as if they were. This homoplasy could increase apparent similarity of populations (Estoup *et al.* 2002). In addition, if relatively few markers are used in the analyses, the frequency of those alleles in the population may have remained stable during isolation and unaffected by forces of evolution. If insufficient markers
are investigated, or the microsatellite markers chosen do not present differentiation where other markers may, the investigation may be prone to type II error, suggesting that there is not a difference when significant differentiation would be observed if analyzed by alternative markers. This may occur through poor marker choice or a large population with relatively stable allele frequencies. Increasing the number of polymorphic markers analyzed can allow characterization of populations with higher resolution. Since each marker has its own individual history, more markers could allow a more comprehensive investigation of drift between two isolated populations.

Another factor that may affect the outcome of microsatellite analyses is complications due to null alleles. Null alleles are alleles that do not amplify in certain individuals. This may be due to polymorphism at the primer site or failure of the primer to anneal to the DNA template strand during PCR. The failure of amplification is a problem because if one allele is not amplified, the other may be preferentially amplified during PCR, leading to an increased abundance of the single amplified fragment. The absence of the non-amplified fragment leads to the characterization of the individual as homozygous when in fact the individual is heterozygous (Dakin & Avise 2004). The implication of apparent homozygosity is that it alters characterization of the population and ultimately affects measures of genetic distance between populations.

The use of a relatively underutilized technique may overcome some of these problems and provide a clearer depiction of population differentiation of bay scallops in North Carolina. Amplified Fragment Length Polymorphism (AFLP) is the selective PCR amplification of restriction fragments from genomic DNA. The process consists of four phases: (1) restriction digest of genomic DNA by two restriction enzymes, (2) attachment of double stranded restriction
enzyme primer adaptors, (3) preselective amplification using primers that attach to enzyme adaptors and overlap the fragment by one nucleotide, and (4) selective amplification using preselective template-specific primers with two or three overlapping nucleotides. After selective amplification, fragments may be analyzed by electrophoresis without prior knowledge of the nucleotide sequence. This technique combines the power of PCR with the reliability of RFLPs (Vos et al. 1995).

There are several advantages to using AFLP as compared to other techniques. First, it is possible to obtain genetic fingerprints with small amounts of DNA because of the utilization of PCR in the technique. The application of PCR allows the simultaneous amplification of many markers for use in the analysis. Secondly, AFLP works on a variety of organisms. It uses restriction enzymes that are able to recognize and cut genomic DNA, creating fragments of varying sizes. Since AFLP utilizes restriction enzymes and PCR using primers incorporated as synthetic adapters, it is not necessary to have prior knowledge of genomic sequences. When different primer pairs are used, unique fragment fingerprints can be obtained and statistically analyzed across samples to draw conclusions (PE Applied Biosystems 1997). The use of AFLP in this study will allow a powerful technique to assist in evaluating the level of population structure in North Carolina coastal sounds.

Objective

The objective of this study is to determine if there is local bay scallop population structure in North Carolina coastal bays based on AFLP analysis.
Significance

The data from this genetic structure analysis will allow management to determine whether they should focus their effort on several populations or a single population. Although this study will not identify all populations in North Carolina, if there is population structure in North Carolina, this study will provide management agencies with the knowledge that there are indeed separate populations that need to be considered when designing an effective management or restoration plan. If there are multiple populations of bay scallops in North Carolina, more intensive studies can take place investigating the interactions between populations of interest and these studies could eventually allow the identification of source and sink populations.

Management can then better protect source populations so that adults from that population can contribute gametes during a spawn. If the bay scallop population in North Carolina is panmictic, a single statewide management plan can be created to protect the population as a single unit.
METHODS

Sample Collection

In 2007, approximately 50 scallops were collected from each of eight North Carolina sites (Figure 1, Table 1).

Table 1. Scallops collected for genetic analysis and site abbreviations.

<table>
<thead>
<tr>
<th>Site</th>
<th>Site Code</th>
<th>Water Body</th>
<th>Date</th>
<th>Number of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hewletts Creek</td>
<td>HC</td>
<td>Masonboro Inlet</td>
<td>6-4-08</td>
<td>55</td>
</tr>
<tr>
<td>New River Inlet</td>
<td>NR</td>
<td>New River</td>
<td>5-28-08</td>
<td>64</td>
</tr>
<tr>
<td>Hammocks Beach</td>
<td>HB</td>
<td>White Oak River</td>
<td>8-19-08</td>
<td>51</td>
</tr>
<tr>
<td>Western Bogue</td>
<td>WB</td>
<td>Bogue Sound</td>
<td>6-25-08</td>
<td>56</td>
</tr>
<tr>
<td>Eastern Bogue</td>
<td>EB</td>
<td>Bogue Sound</td>
<td>7-28-08</td>
<td>54</td>
</tr>
<tr>
<td>Harkers Island</td>
<td>HI</td>
<td>Back Sound</td>
<td>6-18-09</td>
<td>54</td>
</tr>
<tr>
<td>Bells Island</td>
<td>BI</td>
<td>Core Sound</td>
<td>7-17-08</td>
<td>63</td>
</tr>
<tr>
<td>Ocracoke Island</td>
<td>OI</td>
<td>Pamlico Sound</td>
<td>10-7-08</td>
<td>55</td>
</tr>
<tr>
<td>New York</td>
<td>NY</td>
<td>Long Island Sound</td>
<td>1999</td>
<td>47</td>
</tr>
</tbody>
</table>

Figure 1. Sampling sites. See Table 1 for explanation of the abbreviations.
Scallops were collected haphazardly by hand and put on ice within 20 minutes of collection to prevent degradation of DNA. Once scallops were transported to the analysis facility, they were stored at -20˚C until processed. A sample collected from Long Island Sound in New York during 1999 was used as an “outgroup” in this analysis.

DNA Extraction

DNA extraction was performed on adductor tissue using a PureGene DNA extraction kit (Qiagen, Gaithersburg, MD). Briefly, 200 μl of cell lysis solution and 1.5 μl of 20 mg/ml Proteinase-K were added to 2-4 mg of adductor tissue and allowed to incubate at 55˚C overnight. The following day, samples were removed from the water bath and allowed to cool to room temperature. Seventy μl of protein precipitate solution was added to sample tubes and each sample was vortexed with a VWR mini vortexer on a speed of 7 for 20 seconds and placed on ice. The samples were incubated on ice for 5 minutes and spun in an Eppendorf centrifuge 5424/5415D for 8 minutes at 14,000 rpm. The resulting supernatant was transferred to a new 1.5 ml tube and 200 μl of 100% isopropanol was added. Sample tubes were inverted 50 times and spun for 8 minutes at 14,000 rpm. The supernatant was discarded and the pellet was incubated at room temperature for 5 minutes. Two hundred μl of 70% ethanol was added to the tubes and allowed to soak the DNA for 5 additional minutes. The tubes were then spun for 8 minutes at 14,000 rpm. The supernatant was discarded and the pellet was dried overnight at room temperature. The following day, the pellet was rehydrated in 35 μl of sterile dH2O for 3 hours and frozen for future analysis.
AFLP Analysis

DNA concentration of the extracts was quantified as the average of triplicate measures with a NanoDrop ND-1000 spectrophotometer and diluted with sterile dH₂O so that the concentration of DNA was in the range of 200-1000 ng/μl, the concentration recommended by PE Applied Biosystems (ABI, Foster City, CA).

AFLP analysis was conducted using a modified version of the ABI AFLP Large Genome Plant Mapping Kit protocol. Briefly, enough Enzyme Master Mix was prepared (0.06 μl 10x T4 Ligase Buffer, 0.06 μl 0.5M NaCl, 0.025 μl 1mg/ml BSA, 0.06 μl 50,000 U/ml MseI, 0.06 μl 100,000 U/ml EcoRI, 0.025 μl 400,000 U/ml T4 Ligase) for each reaction. The master mix was mixed gently and stored on ice for use in the restriction-ligation reaction. Primer adapters were denatured as per the protocol and stored at room temperature for the restriction ligation reaction.

Restriction ligation master mix was prepared (0.5 μl 10x T4 DNA ligase buffer w/ ATP, 0.5 μl 0.5M NaCl, 0.25 μl 1.0 mg/ml BSA, 0.5 μl kit supplied MseI adaptor, 0.5 μl kit supplied EcoRI adaptor, 0.5 μl Enzyme master mix, 0.5 μg restricted DNA in 2.75 μl sterile H₂O). The mixture was inverted 10 times, spun momentarily in a microcentrifuge, and incubated overnight at room temperature. The following day, 98.75 μl 0.1 M TE buffer (20ml 0.1 M Tris HCl, 20ul 0.5 M EDTA, filled to 100ml with water) was added to each reaction to dilute the solution to the proper concentration for preselective amplification. The mixture was finger vortexed, spun momentarily in a microcentrifuge, and stored at 5°C.

The diluted restriction-ligation reactions (2 μl) were mixed with 0.25 μl each of the preselective primers and 7.5 μl of AFLP core mix then amplified on a PTC-100 or PTC-200. Amplification conditions were composed of 2 min at 72 °C followed by 20 cycles of 20 sec at 94
°C, 30 sec at 56 °C and 2 min at 72 °C, followed by 30 min at 60 °C, and a final cold storage at 5°C.

Preamplification product (10 μl) was diluted with 190μl of 0.1 M TE buffer, finger vortexed, and spun momentarily in a microcentrifuge. Selective amplification reactions (1.5 μl dilute preselective amplification product, 0.5 μl of 5μM MseI [Primer- Cxx], 0.5 μl of 1μM EcoRI [Dye- Primer- Axx], 7.5 μl of AFLP core mix) were prepared using one of the following 6 primer pair combinations: A: EcoRI primer-ACG with MseI primer-CAT, B: EcoRI primer-ACT with MseI primer-CAG, C: EcoRI primer -AAC with MseI primer-CTT, G: EcoRI primer-ACG with MseI primer-CTT, H: EcoRI primer-ACT with MseI primer-CAT, I: EcoRI primer-AAC with MseI primer-CAG. Amplifications were performed with the following temperature profile: 2 min at 94 °C, 10x(20 sec at 94 °C, 30 sec at 66 °C[-1°C each cycle], 2 min at 72 °C), 20x[20 sec at 94 °C, 30 sec at 56 °C, 2 min at 72 °C], then 30 min at 60°C.

Selective amplification products (1 μl) were mixed with 8 μl of formamide-ROX size standard loading buffer (37.5 μl GeneScan -500 ROX size standard in 1 ml Hidi). Primer pairs tagged with distinct dyes were mixed (1 μl A+ 1 μl B+ 2 μl C, 1μl G+ 1μl H+ 1.5 μl I), denatured and placed on ice until loaded on an ABI 3130x/Genetic analyzer.

Fragment Detection and Signature Evaluation

Genetic analyzer AFLP fragment data profiles were imported into GENEMAPPER. Analysis parameters were specified to include a modified factory provided ROX size standard setting of (GS500(-35-250)). The GENEMAPPER analysis method was configured to capture the dye of interest based on the primer pair investigated. Panels were generated automatically.
with a 1 base pair bin width from samples incorporating all samples to prepare proper binning. Peak heights were normalized to the sum of signal, common alleles were not deleted, and binary allele calling was used. The peak quality tab used an advanced peak detection algorithm and peak detection level was set to background Relative Florescence Units (RFU), a measure of emission intensity. All other options were left as default. The advanced scoring setting was used in order to allow detection of fragments above background level as determined by eye as the highest background level for the noisiest sample for each primer pair.

In order to isolate “real” peaks from background noise or artificial peaks due to pull-up or double peaking, each primer pair was analyzed with the threshold setting set just above background level and was increased until the automatic bins registered 100 ± 20 bins per primer pair set to capture a manageable number of markers. Any samples that did not have a size quality RFU high enough for reliable detection of fragment sizes could not be analyzed and were discarded. After the removal of these poor quality samples, the remaining data was re-analyzed to ensure correct automatic panel generation. The resulting panel was exported to the panel manager for each primer pair.

Samples were then analyzed with the GENEMAPPER analysis method incorporating the exported panel settings for each primer pair. The presence/absence threshold level was set to just above background level, determined as described above. Detection threshold level was lowered from the previous analysis to allow detection of peaks that produced fluorescence at a given fragment size, but did not reach the stringent detection threshold used to establish bins. In this fashion, bins constructed based on highly fluorescent peaks, indicative of non-artifact peaks, would limit the presence of artifact fragments, but would reduce the amount of stringency within bins and reduce type II error due to poor peak detection at high thresholds. Peak detection level
was set to 10 RFUs, lower than any background level. Primer pair sets were analyzed throughout their reliable fragment calling size ranges (A= 71-500 bp, B= 75-500 bp, C= 80-500 bp, G= 50-250 bp, H= 50-220, I= 50-251 bp).

Resulting binary peak presence/absence data was exported into Microsoft Excel. Binary data for each primer pair were aligned for each sample. Only samples that had bands present for all six primer pairs were used in the analysis. Primer pair databases were combined, exported, and formatted into text files for statistical analyses.

Statistical Analysis

The average number of fragments per primer pair and average number of fragments per site were calculated to ensure that neither one site nor one primer pair had more of an impact on the study than any other site or primer pair. In addition, a histogram of the number of AFLPs per individual was constructed to make sure the data was unimodal. This was performed to ensure that no subset of individuals at any specific primer pair had a drastically different number of AFLPs relative to other individuals used in the study. Unimodality was also ensured across all fragments for all individuals to so that no subset of individuals was confounding the analysis. To describe how different each population was from one another, a measure of the number of average fragment differences was calculated using a parsimony distance matrix (total number of fragment differences between individuals) with PAUP V.4 (Swofford 2002). The average fragment differences among all individuals, within populations, and among populations were calculated in MEGA 4 (Tamura et al. 2007).
Population differentiation was evaluated by multiple statistical approaches. An analysis of molecular variance (AMOVA) (Excoffier et al. 1992) was used to assess the proportion of variance attributed to population structure based on geographic regions (Arlequin, Excoffier et al. 2005). Pairwise $\Phi_{ST}$ were estimated to measure the amount of genetic differentiation between each combination of populations against the null hypothesis of homogeneity. These $\Phi_{ST}$ values were calculated by comparing the number of different AFLPs between individuals in each population.

Assignment tests were performed in AFLPOP 1.1 (Duchesne & Bernatchez 2002). In this analysis, each individual’s fingerprint was allocated to source populations based on the differences in the component fragment frequencies in each source population. Population fragment frequencies were calculated after each focal individual was removed (“leave-one-out” procedure, Waser & Strobeck 1998), eliminating the bias of assigning individuals to populations to which they have contributed baseline data. Individuals were then assigned to a population based on the likelihood that the removed signature matched the frequency profile of that population better than any other. Accuracy was assessed by determining the number of scallops that were correctly assigned to their home populations.

Bayesian clustering analysis of individuals was conducted using STRUCTURE version 2.2 (Pritchard & Wen 2003). Bayesian clustering groups similar individuals together based on shared AFLPs and can be used to determine whether there is genetic structure, to show the proportion of individuals assigned from each location to each cluster, and to show the individual's strength of clustering to each group. Using a modification to the method developed by Falush et al. (2007), AFLP data was formatted for STRUCTURE as duplicated binary data and analyzed as a diploid data set because analyses performed as co-dominant markers decreases
genotypic ambiguity and returns a more accurate results. The “no admixture model” and the “correlated allele frequency” settings were selected as suggested for populations that may have a small amount of differentiation (Pritchard & Wen 2003). The program was run for a range of groups (K) from 1-6 with three independent runs performed for each K (60,000 burn-in, 100,000 run time). Burn-in and run time were determined empirically by comparing multiple replicate runs for consistency. Data was then imported into STRUCTURE HARVESTER (Earl 2009) to graphically display the STRUCTURE output and display the average –log likelihood values for each K. “True K” (Falush et al. 2007), a quantity based on the second order rate of change of the likelihood function with respect to K, was determined as described by Evanno et al. (2005).

Once the true K was determined, 10 iterations of STRUCTURE analysis were run at the true K with a burn-in of 60,000 and a 100,000 iteration run. The output was compressed and imported to STUCTURE HARVESTER (Earl 2009). The CLUMPP infile created by STRUCTURE HARVESTER was imported into CLUMPP to cluster independent runs of the STRUCTURE output so that multiple replicates could be averaged to decrease random error created by the Markov chain Monte Carlo simulation process (Jakobsson & Rosenberg 2007). Triplicate runs of both regions were run with “full search” option selected. For North Carolina scallops, the “true K” was then run 10 additional times and CLUMPP was run with a Greedy search option and random input orders with 1,000 permutations. Output files from CLUMPP were then formatted for use in the program DISTRUCT to graphically represent the associated assignments to each group (Rosenberg 2004).
RESULTS

Descriptive Statistics

A total of 302 individual scallops were used in the genetic analyses. The number of individuals per population ranged from 18-48 with an average of 34 individuals. Sample numbers per site used for statistical analysis were as follows: Hewletts Creek (HC) n=30, New River (NR) n=37, Hammocks Beach (HB) n=44, Western Bogue (WB) n=34, Eastern Bogue (EB) n=28, Harkers Island (HI) n=28, Bells Island (BI) n=48, Ocracoke Island (OI) n=18, New York (NY) n=35. Descriptive statistics of the number of fragments for each primer pair are provided in Table 2. There were an average of 79 bins per primer pair and no two individuals shared the same combination of AFLPs. Four hundred forty seven variable markers were used in the analysis. Of the 447 variable markers, 13 markers were singletons and 434 markers were polymorphic.

Table 2. AFLP fragment summary statistics by primer pair.

<table>
<thead>
<tr>
<th>Primer Pair</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>All Primer Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of fragment bins</td>
<td>84</td>
<td>81</td>
<td>59</td>
<td>74</td>
<td>86</td>
<td>85</td>
<td>469</td>
</tr>
<tr>
<td>Maximum number of fragments present</td>
<td>40</td>
<td>29</td>
<td>30</td>
<td>33</td>
<td>50</td>
<td>46</td>
<td>200</td>
</tr>
<tr>
<td>Min number of fragments present</td>
<td>20</td>
<td>14</td>
<td>12</td>
<td>19</td>
<td>24</td>
<td>26</td>
<td>152</td>
</tr>
<tr>
<td>Average number of fragments present</td>
<td>31</td>
<td>23</td>
<td>22</td>
<td>25</td>
<td>39</td>
<td>36</td>
<td>176</td>
</tr>
<tr>
<td>Standard deviation of fragments present</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>9</td>
</tr>
</tbody>
</table>

The average number of fragments per site for each primer pair is depicted in Figure 2. The number of present fragments of each primer pair was similar from site to site. The fragment
frequency distributions for each primer pair can be seen in Figure 3. The histogram shows a normal distribution for each primer pair. The number of fragments for each individual was summed and plotted to ensure a normal distribution of total fragments among individuals (Figure 4). The average number of total fragments \( \pm \text{SD} \) for each site was as follows: BI yielded 173\( \pm \)9 fragments, EB yielded 173\( \pm \)6 fragments, HB yielded 176\( \pm \)9 fragments, HC yielded 175\( \pm \)9 fragments, HI yielded 175\( \pm \)10 fragments, NR yielded 175\( \pm \)7 fragments, OI yielded 177\( \pm \)7 fragments, WB yielded 178\( \pm \)8 fragments, and NY yielded 182\( \pm \)8 fragments. A fragment distance matrix produced by PHYLIP and subsequent matrix analysis by MEGA showed that the individuals differed by an average difference of 101 fragments.

Figure 2. The average number of fragments per site for each primer pair (\( \pm \text{SD} \)) show similar fragment numbers at each site across each primer pair. Sky blue fill represents fragments of primer pair A, teal fill represents fragments of primer pair B, light blue represents fragments of primer pair C, blue represents fragments of primer pair G, navy blue fill represents fragments of primer pair H, dark blue fill with black outline represents fragments of primer pair I.
Figure 3. Frequency distributions of the average number of fragments for each primer pair show unimodal distributions of the number of individuals for each primer pair. Dark blue fill represents fragments of primer pair A, navy blue represents fragments of primer pair B, blue represents fragments of primer pair C, light blue represents fragments of primer pair G, teal fill represents fragments of primer pair H, sky blue fill represents fragments of primer pair I.

Figure 4. Frequency distribution of total fragments across all primer pairs show a unimodal distribution.
Population Comparisons between New York and North Carolina

Cluster Analysis

Clustering analysis by STRUCTURE, a program that does not assume any structure \textit{a priori}, did not show a strong peak in $-\log$ likelihood values at any particular number of clusters (Figure 5). Plots of $L(K)$ did show a weak peak $K=5$ so the Evanno method was implemented to find the change of the $2^{nd}$ order rate of $K$ to provide a better estimate of the “true” $K$. Implementation of the Evanno method suggested that the data did sort best into five clusters (Figure 6). Three independent runs of STRUCTURE at $K=5$ were combined using CLUMPP to incorporate multiple runs. The resulting values were run in DISTRUCT to convert the data to graphical form (Figure 7). The graphical output at $K=5$ showed strong individual clustering $Q$-values and showed a strong difference between scallops from North Carolina and from New York with an average $Q$ of 0.985 for scallops that were allocated to New York. The $Q$-statistic in STRUCTURE is the estimated membership coefficient of each individual or, in the case of population Qs, the average coefficient of all individuals in a sample population in each cluster. Individual and population $Q$-statistics are shown in Figure 7. STRUCTURE’s priorloc models and admixture models were investigated to determine if there were any noticeable changes in population or individual Qs, but no such changes were observed (Pritchard \textit{et al.} 2000).


Figure 5. STRUCTURE output of triplicate runs from K=1-6 including New York do not show a pronounced peak. STRUCTURE results showed that the first order rate of change of the likelihood distribution with increasing K did not produce a strong peak at a “true” K. Individual runs are denoted by closed circles, averages are denoted by open squares flanked by standard deviation error bars.

Figure 6. Graphical description of implementation of the Evanno method including New York. Plot of L(K) for K=1-6 (left). The Evanno method was implemented to determine the true K to be used in this analysis by investigating the second order rate of change of the likelihood distribution of the data (middle). The change of the second order rate of change of the likelihood distribution showed a pronounced peak at K=5 (right).
Figure 7. Q-statistics including New York at K=5. Above: Average population Qs at K=5 including New York. Each color reflects each of five clusters to which individual were allocated. The height of the color corresponds to the proportions of individual in each population allocated to that cluster. The width of each site corresponds to the relative number of individuals analyzed at each site. Below: Individual Qs at K=5 including New York. Each color represents each of the five clusters to which each individual could be allocated. The height of the color corresponds to the proportion of clusters for which each individual was allocated. The width of each site corresponds to the relative number of individuals analyzed at each site.

Analysis of molecular variance

AMOVA analysis groups were constructed to explain the most variation relative to other grouping strategies. With all data, the AMOVA grouping strategy that explained the most variation was \( \{\text{HC, NR, HB, WB, EB, HI, BI, OI}\} \{\text{NY}\} \) (Table 3). \( \Phi_{ST} \) was estimated to be 0.06249, a value suggestive of subtle genetic structure (P<0.0001). Among populations within group variation was not statistically significantly less than the value expected under the null hypothesis of homogeneity (\( \Phi_{SC}=0.01321, P=0.12317 \)). The proportion of variation among groups was statistically significantly different from the value expected under homogeneity (\( \Phi_{CT}=0.04994, P<0.0001 \)).
Table 3. Arlequin AMOVA results with groupings of {HC, NR, HB, WB, EB, HI, BI, OI} {NY}.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among groups</td>
<td>1</td>
<td>235.923</td>
<td>2.63668</td>
<td>4.99</td>
</tr>
<tr>
<td>Among populations within</td>
<td>7</td>
<td>499.650</td>
<td>0.66258</td>
<td>1.26</td>
</tr>
<tr>
<td>groups within populations</td>
<td>293</td>
<td>14501.706</td>
<td>49.49388</td>
<td>93.75</td>
</tr>
<tr>
<td>Within populations</td>
<td>301</td>
<td>15237.278</td>
<td>52.79314</td>
<td></td>
</tr>
</tbody>
</table>

Pairwise $\Phi_{ST}$s show that New York was significantly different from every North Carolina population ($P \leq 0.00001$, Table 4). Twenty seven of 36 population comparisons were significantly different from each other. Nineteen of the 30 significantly different population comparisons were significant at $P < 0.0001$.

Table 4. Pairwise $\Phi_{ST}$s and significance matrix from Arlequin including New York after 10100 permutations. $\Phi_{ST}$s are displayed below the division line. Significance after a sequential Bonferroni adjustment ($\alpha = 0.05$) are displayed above the division. “*” denotes significant comparison. “-” denotes comparisons that were not significant.

<table>
<thead>
<tr>
<th></th>
<th>HC</th>
<th>NR</th>
<th>HB</th>
<th>WB</th>
<th>EB</th>
<th>HI</th>
<th>BI</th>
<th>OI</th>
<th>NY</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td>0.01433</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HB</td>
<td>0.01346</td>
<td>0.00873</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>-</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WB</td>
<td>0.02013</td>
<td>0.01306</td>
<td>0.01472</td>
<td>-</td>
<td>-</td>
<td>*</td>
<td>-</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>EB</td>
<td>0.02097</td>
<td>0.02053</td>
<td>0.01969</td>
<td>0.00798</td>
<td>-</td>
<td>*</td>
<td>-</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>HI</td>
<td>0.02017</td>
<td>0.01833</td>
<td>0.01573</td>
<td>0.00541</td>
<td>0.00381</td>
<td>*</td>
<td>-</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>BI</td>
<td>0.01591</td>
<td>0.00441</td>
<td>0.00202</td>
<td>0.01024</td>
<td>0.01995</td>
<td>0.01433</td>
<td>-</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>OI</td>
<td>0.02462</td>
<td>0.01866</td>
<td>0.01455</td>
<td>0.00678</td>
<td>0.00111</td>
<td>0.0047</td>
<td>0.01097</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>NY</td>
<td>0.07019</td>
<td>0.06255</td>
<td>0.06719</td>
<td>0.05704</td>
<td>0.07039</td>
<td>0.06472</td>
<td>0.06599</td>
<td>0.06618</td>
<td></td>
</tr>
</tbody>
</table>
Assignment test

When all sites were assessed as separate populations in a reassignment test, 100% individuals tested from North Carolina were assigned to North Carolina and 94% of individuals from New York were assigned back to New York, suggesting strong differentiation (results not shown).

Population Comparisons within North Carolina

Cluster Analysis

A STRUCTURE clustering analysis of North Carolina populations suggested that the data did not group, as evidenced by the absence of a clear peak, at a particular K (Figure 8). Log likelihood increased as K was increased and plateaued. At K=6 one of the triplicate runs was showed a strong negative value. To investigate the legitimacy of the data point, seven additional runs at K=6 and two runs at K=7 were performed and resulted in – log likelihood values in agreement with the plateau pattern. Thus, for the three replicates needed for the Evanno method at K=6, data was used that followed the pattern of the majority of the data. Plots of L(K) did show a small peak at K=4 but the Evanno method was implemented to find the change of the 2nd order rate of K (Figure 9). The Evanno method produced a clear peak at K=4, supporting the conclusion from the original STRUCTURE runs. Thirteen independent runs of STRUCTURE at K=4 were combined using CLUMPP to incorporate repetitive runs (Table 5). The resulting graphical output at K=4 showed strong individual clustering Q-values (Figure 10, below), but did not show strong distinctions among population Qs (Figure 10, above). There does appear to be modest differentiation of the central populations (Western Bogue, Eastern Bogue, Harkers Island) relative to those to the North and South.
Figure 8. STRUCTURE output of North Carolina triplicate runs from K=1-6. STRUCTURE results showed that the first order rate of change of the likelihood distribution with increasing K did not produce a clear peak at a “true” K.

Figure 9. Graphical description of the Evanno method for North Carolina populations. Plot of the North Carolina L(K) for K=1-6 (left). The Evanno method was implemented to determine the true K to be used in this analysis by investigating the second order rate of change of the likelihood distribution of the data (middle). The change of the second order rate of change of the likelihood distribution showed a pronounced peak at K=4 (right).
Table 5. STRUCTURE output of 13 replicate North Carolina runs at K=4.

<table>
<thead>
<tr>
<th>K</th>
<th>Est. Ln Prob. of Data</th>
<th>Mean Value of Ln likelihood</th>
<th>Variance of Ln likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-60795.2</td>
<td>-57549.5</td>
<td>6491.4</td>
</tr>
<tr>
<td>4</td>
<td>-60758.7</td>
<td>-57550.3</td>
<td>6416.8</td>
</tr>
<tr>
<td>4</td>
<td>-60728.4</td>
<td>-57549.5</td>
<td>6357.7</td>
</tr>
<tr>
<td>4</td>
<td>-60795.3</td>
<td>-57549.6</td>
<td>6491.4</td>
</tr>
<tr>
<td>4</td>
<td>-60827.7</td>
<td>-57554.6</td>
<td>6546.2</td>
</tr>
<tr>
<td>4</td>
<td>-60889</td>
<td>-57554.1</td>
<td>6669.8</td>
</tr>
<tr>
<td>4</td>
<td>-60814.5</td>
<td>-57550.1</td>
<td>6528.7</td>
</tr>
<tr>
<td>4</td>
<td>-60875.9</td>
<td>-57600.5</td>
<td>6550.8</td>
</tr>
<tr>
<td>4</td>
<td>-60805</td>
<td>-57545.7</td>
<td>6518.6</td>
</tr>
<tr>
<td>4</td>
<td>-60761.8</td>
<td>-57549.5</td>
<td>6424.7</td>
</tr>
<tr>
<td>4</td>
<td>-60893.1</td>
<td>-57550.7</td>
<td>6684.8</td>
</tr>
<tr>
<td>4</td>
<td>-60759.6</td>
<td>-57549</td>
<td>6421.1</td>
</tr>
<tr>
<td>4</td>
<td>-60831.8</td>
<td>-57549.7</td>
<td>6564.2</td>
</tr>
</tbody>
</table>

Figure 10. Q-statistics of North Carolina populations at K=4. Above: Average population Qs at K=4 for North Carolina. Each color reflects each of four clusters to which individual were allocated. The height of the color corresponds to the proportions of individual in each population allocated to that cluster. The width of each site corresponds to the relative number of individuals analyzed at each site. Below: Individual Qs at K=4 for North Carolina. Each color represents each of the four clusters to which each individual could be allocated. The height of the color corresponds to the proportion of clusters for which each individual was allocated. The width of each site corresponds to the relative number of individuals analyzed at each site.
Analysis of molecular variance

AMOVA analysis of only the North Carolina populations after revealed subtle differentiation when the samples were grouped by the pattern inferred by STRUCTURE ({HC, NR, HB}{WB,EB, HI}{BI, OC})(Table 6). $\Phi_{ST}$ was estimated to be 0.01386, a value suggestive of very subtle, but significant genetic structure (P<0.0001). Among populations within group variation was not statistically significantly different from the value expected under the null model of homogeneity ($\Phi_{SC}=0.00951$, P=0.08895). The proportion of variation among groups was statistically significantly different from the value expected under homogeneity ($\Phi_{CT}=0.00439$, P<0.0001).

Table 6. AMOVA results of North Carolina populations grouped by geographic regions. {HC, NR, HB}{WB, EB, HI}{BI, OI}. See Table 1 for abbreviations.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among Groups</td>
<td>2</td>
<td>173.394</td>
<td>0.22263</td>
<td>0.44</td>
</tr>
<tr>
<td>Among populations w/in groups</td>
<td>5</td>
<td>326.256</td>
<td>0.47996</td>
<td>0.95</td>
</tr>
<tr>
<td>Within populations</td>
<td>259</td>
<td>12946.163</td>
<td>49.98518</td>
<td>98.61</td>
</tr>
<tr>
<td>Total</td>
<td>266</td>
<td>13445.813</td>
<td>50.68778</td>
<td></td>
</tr>
</tbody>
</table>

Population differentiation within North Carolina was estimated with pairwise $\Phi_{STS}$.

Pairwise $\Phi_{ST}$ results after sequential Bonferroni adjustment of critical values showed significant differentiation between 20 of the 28 North Carolina population comparisons investigated (P<0.001) (Table not shown). Contrary to the analysis including New York, the change in critical values from the Bonferroni adjustment differentiated the Ocracoke population from the Bells Island population. All other relationships remained consistent with the analysis including
New York. Hewletts Creek samples were significantly different from all other populations in North Carolina. Ocracoke appeared to be similar to more southern regions including Western Bogue (P= 0.05019), Eastern Bogue (P= 0.38531), and Harkers Island (P= 0.14028). Western Bogue and Eastern Bogue were not statistically different (P= 0.00891) and Eastern Bogue was similar to Harkers Island (P=0.13504). There was also a relatively low $\Phi_{ST}$ between scallops from Bells Island and those from Hammocks Beach ($\Phi_{ST} = 0.00202$, P=0.16731). All other populations were significantly different from each other (P<0.005).

Assignment test

The assignment test assigned twenty-six percent of the individuals back to their source population, indicating a lack of genetic distinctiveness among populations (Table 7). When populations, however, are grouped into regions as previously described (\{HC, NR, HB\} \{WB,EB,HI\} \{BI, OC\}), 48.7% of individuals were correctly assigned back to their source region (Table 7). Correct regional assignment increases to 58% on average for the central three populations of WB, EB and HI reflecting the subtle differentiation of these populations (Table 7, Figure 11) relative to the others.
Table 7. AFLPOP reassignment test (leave-one-out) distribution of North Carolina populations. Clustering was performed to group similar populations that were often assigned incorrectly to each other. Three clusters were examined. One consisted of Hewletts Creek, New River, Hammocks Beach (red). The second consisted of Western Bogue, Eastern Bogue, and Harkers Island (yellow). The third consisted of Bells Island and Ocracoke Island (green).

<table>
<thead>
<tr>
<th>Allocated to</th>
<th>HC</th>
<th>NR</th>
<th>HB</th>
<th>WB</th>
<th>EB</th>
<th>HI</th>
<th>BI</th>
<th>OI</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>13</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>NR</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>HB</td>
<td>5</td>
<td>7</td>
<td>12</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>WB</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>11</td>
<td>7</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>EB</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>HI</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>BI</td>
<td>6</td>
<td>12</td>
<td>17</td>
<td>8</td>
<td>2</td>
<td>7</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>OI</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 11. Regional allocation of assignment tests within North Carolina. Pie charts indicate the proportion of each population that were assigned to each region (south=red, central=yellow north=green).
DISCUSSION

Samples from New York were compared to samples from North Carolina in order to distinguish the amount of variation between large geographic regions based on a variety of methods. Global $\Phi_{ST}$ (including New York) was estimated at 0.0627, suggestive of subtle, yet significant, population structure. Scallops from New York and North Carolina were not found to be differentiated when analyzed using microsatellites (Hemond 2006). The results of the STRUCTURE analysis ($K=5$) based on AFLP data showed strong differentiation between scallops from North Carolina and New York in contrast to the findings from microsatellite analysis. In addition, a re-assignment test based on the AFLP data correctly allocated 99.3% of scallops to their correct source state which was substantially higher than the less than 50% based on microsatellites observed by Hemond (2006).

The finding of population structure in scallops in this study was not completely attributable to the distinctiveness of scallops from New York. The $\Phi_{ST}$ estimate based on the analysis of only North Carolina populations was significant, in agreement with the study based on mtDNA sequence analysis performed by Marko and Barr (2007). Both studies suggest that there is significant genetic population structure within North Carolina.

A subject of particular interest from both previous studies was differentiation of populations across Beaufort Inlet. Hemond (2006) did not find significant differentiation between across the inlet. Marko and Barr (2007), on the other hand, did observe subtle, but significant differentiation across the inlet and suggested that genetic differentiation was established due to river outflow through the inlet preventing the exchange of larvae or scallops between these two basins. The pairwise $\Phi_{ST}$ between East Bogue and Harkers Island in this
study did not detect significant differentiation between populations across that same inlet (Table 4).

Although not directly comparable, similar estimates of genetic differentiation have been observed in other broadcasting species with similar life history patterns across similar spatial scales to the 200 km range in this study. Zhan et al. (2009) observed F_{ST}s ranging from 0.0013-0.0602 derived from nine microsatellite markers in Zhikong scallops, *Chlamys farreri*, from Northern China. The scallops analyzed were from six populations. Two of the populations investigated were located just 92 km from one another, but showed significant differentiation (F_{ST}=0.0422). Zhao et al. (2009) observed an F_{ST}=0.0264 between two wild populations of Zhikong scallops in Northern Japan with 293 AFLP markers across less than 450 km. Kenchington et al. (2006) observed a F_{ST}=0.011 with respect to sea scallops, *Placopecten magellanicus*, along the East Coast of North America using six microsatellite markers. Within a 300 km range across Georges Bank, the study observed a θ of 0.046. Banks et al. 2007 also observed a small, but significant level of population differentiation in sea urchins, *Centrostephanus rodgersii*, across a range of 3,000 km (F_{ST}=0.008, P<0.001). A general pattern of increasing differentiation has been observed with increasing geographic distance, and is termed “isolation by distance” or IBD (Wright 1943). IBD has been noted in other broadcast spawning species such as the eastern oyster (*Crassostrea virginica*) in Chesapeake Bay which displays significant differentiation among populations separated by 1-345 km (Rose et al. 2006). Contrastingly, Banks et al. (2007) observed that neighboring populations were more differentiated than geographically distant populations. Of sites less than eight km apart, mean θ was 0.025 (Banks et al. 2007).
Although not directly tested, isolation by distance is not apparent in this study because $\Phi_{ST}$ of broadly separated populations are not significantly different from the null hypothesis of homogeneity. Estimates of $\Phi_{ST}$ appear to suggest regional differentiation, rather than differentiation of individual populations, as do of the results of the reassignment test. Scallops that were not assigned to their source population were more likely to be “miss-assigned” to a population within the same region (Figure 11).

A similar genetic pattern is found in the results from the clustering analysis. STRUCTURE results suggest that individuals have high $Q$-values, but the clusters are widely distributed geographically (Figure 10). The region of West Bogue, East Bogue, and Harkers Island contained the highest frequency of one cluster (Figure 10, yellow) relative to populations to the south or north. The reason for such stark differentiation between adjacent sites such as Hammocks Beach and Western Bogue or Harkers Island and Bells Island is unclear. The distance between Hammocks Beach and Western Bogue is just 9.5 km and the two have a $\Phi_{ST}$ of 0.01472. The distance between Harkers Island and Bells Island is 8.0 km and the two have a significant pairwise $\Phi_{ST}$ of 0.01433. Since geographically proximal populations were significantly differentiated from one another, patterns suggestive of isolation by distance were not further investigated because some other mechanism of differentiation must be occurring.

The unusual patterns of clustering could be influenced by movement of scallops for research purposes and restoration efforts in Back and Bogue Sound areas during the 1990’s. Peterson and Summerson (1992) introduced 7,269 scallops to Middle Marsh (Back Sound) and 7,418 to Burders Channel in Western Bogue Sound. For several years following their study, Peterson et al. (1996) transported 135,000 scallops in 1992, 100,000 in 1993, and 150,000 in 1994 from Back Sound to Western Bogue. The human mediated movement of scallops around
the area may have augmented the genetic signatures of recipient populations (West Bogue) to reflect the donor populations (Back Sound). This could help to explain why West Bogue, East Bogue, and Harkers Island appear to be more similar to each other than to other populations in this study.

Another reason for this unusual pattern of genetic differentiation could be a response to water current patterns in the area. Water current patterns have been shown to influence genetic differentiation of populations in several studies (Zahn et al. 2009, Saunders et al. 1986). One study in North Carolina described water current patterns in Onslow Bay by documenting the onset of North Carolina’s first red tide event in 1987 (Tester et al. 1991). The red tide was first noted on the ocean side of Bogue Banks and soon after inoculated Bogue, Beaufort, and Bardens inlet; however, the red tide did not arrive in Core sound until a week after the red tide had arrived in the other sounds. Once established, cell counts in Core Sound never reached the concentration of any of the other sounds and the red tide receded from Core Sound six weeks before any of the other sounds (Tester et al. 1991). The reason that the red tide in Core Sound never reached the concentrations of the other sounds may be due to flushing of Core Sound through Lookout Bight. If this indeed is the water current pattern of the area, pelagic larvae in Core Sound may be flushed out of Cape Lookout, circulated in the South Atlantic Bight and “inoculated” back into the coastal bays with southern surface water transport, similar to the inoculation of the red tide.

A third possible explanation of the pattern could be due to the chaotic patchiness that can occur with a boom or bust species like the bay scallop. Since bay scallops are broad cast spawners and only live from 12-18 months, populations could express temporal and spatial variability depending on the spawning success of each cohort. Therefore, the pattern seen in this study could theoretically be different from the pattern seen the following year. Thus to get a
more comprehensive understanding of the stability of population structure, temporal studies should be undertaken so that effective management strategies can be implemented.
CONCLUSION

Genetically differentiated populations have been recognized as important units of fisheries management to delineate management units and maintain sustainable fisheries (Kenchington et al. 2006, Banks et al. 2007). Delineation of these units is even more critical for commercial species such as bay scallops. The genetic differentiation observed in this study suggests that bay scallop populations in North Carolina are subtly, yet significantly, structured throughout the state. The dynamics behind the mechanism of differentiation is unknown. Therefore, until further studies are performed to explore the interactions between geographically proximal, yet genetically differentiated populations, management agencies should be mindful of the genetic differences across water basins when contemplating hatchery production or transplantation of bay scallops. Implications of this study also support recommendations made by Peterson & Summerson (1992) to close the fishery so that the reproductive potential of the species can be maximized the interactions among the scallop populations are better understood.


of chum salmon *Oncorhynchus keta* around the Pacific Rim’, *Journal of Fish Biology* vol. 73, pp 1256-1266.
