DEVELOPMENT OF MICROSATELLITES FOR THE BAY SCALLOP, *ARGOPECTEN IRRADIANS* (LAMARCK), WITH APPLICATION TO EVALUATING RESTORATION

Elizabeth M. Hemond

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

Department of Biology and Marine Biology

University of North Carolina Wilmington

2006

Approved by

Advisory Committee

Frederick Scharf Michael McCartney

Ami E. Wilbur
Chair

Accepted by

Dean, Graduate School

.

TABLE OF CONTENTS

ABSTRACT

Due to the threats of pollution, overfishing, and harmful algal blooms, populations of the bay scallop have become jeopardized throughout this species range. Proactive restoration efforts have been undertaken in Florida, North Carolina and New York with varying degrees of success. However, the interpretation of the impacts of restoration activities are complicated by a lack of direct assessments. The objective of this study was to evaluate the potential of microsatellite markers and multi-locus assignments in genetic assessment of bay scallop restoration.

Five-hundred and thirteen clones from a genomic library were sequenced and yielded 93 loci, of which five were developed for use in this project. These five loci exhibited relatively high variation (8-25 alleles/locus) and adhered to expectations of Mendelian inheritance and Hardy-Weinberg equilibrium. The parents and offspring of a hatchery spawn were genotyped with the microsatellites and revealed a loss of allelic diversity in the F_1 generation. Parentage analysis indicated highly biased contribution with over 87% of the offspring whose parentage was identified being attributable to a single pair of scallops.

Samples of three regional populations of scallops (Florida, North Carolina and New York) were genotyped using nine loci and exhibited significant genic differentiation. The highest level of differentiation was found to occur between Florida and the Atlantic samples ($F_{ST} \geq 0.1137$). Assignment of scallops to regions indicated high power of assignment between Florida and Atlantic populations, but low power between New York and North Carolina populations. In addition, Bayesian clustering indicated population structure between Florida and the Atlantic samples, but none between New York and North Carolina.

Microsatellite analysis and assignment tests were applied to an assessment of restoration in Pine Island Sound following a 2003 release of 1.5 million larvae from a hatchery spawning of

twelve scallops from Anclote Estuary. Scallops sampled from the wild populations in Pine Island Sound, Anclote Estuary and Steinhatchee (a site north of Anclote Estuary) were used to define potential source populations. Little to no differentiation between these populations was observed, and assignment analyses correctly identified the origin of these scallops less than 50% of the time. Assignment of a sample of post-restoration scallops from Pine Island Sound was inconclusive as the assignment tests were unable to distinguish between potential source populations. The utility of microsatellites and multilocus analysis in restoration depends on moderate differentiation between the restoration stock and the wild population, as well as appropriate genetic characterization of the restoration stock sample.

ACKNOWLEDGEMENTS

I would like first to thank Dr. Ami Wilbur, who has been a phenomenal advisor, somehow always maintaining a sense of humor and compassion despite her busy schedule. I am also indebted to my committee members, Dr. Michael McCartney and Dr. Frederick Scharf, for their insights and comments on my project. And thank you to Dr. D. Wilson Freshwater and the other members of his lab for assistance with the sequencer, without which data collection would have been much more difficult.

Additionally, this project would not have been possible without those involved in the aquaculture and fieldwork restoration in Florida, William Arnold, Jay Leverone and Stephen Geiger at the Florida Fish & Wildlife Research Institute. Nor could I have been involved without funding support from NOAA Saltonstall-Kennedy Project.

Further thanks are extended to Katie Schulzitski for help with the microsatellites, Russ Peterson for keeping me entertained in the lab and the occasional work-related assistance, Kristi Sommer for reliably being up for coffee breaks and conversations about genetics and pets, and my friends at the Center for Marine Science.

Finally to my family for their constant support and to Raşit Bilgin for inspiring me to work with conservation genetics and never ceasing to amaze me with his talents and generosity.

DEDICATION

This body of work is dedicated to my grandmothers,

Eva A. Wright and Frances F. Hemond.

LIST OF TABLES

LIST OF FIGURES

BACKGROUND

Bay scallops have historically played an important role as both a commercial and recreational fishery along the Atlantic and Gulf of Mexico coasts of North America. However, over the past few decades depletion of stocks throughout this range has been attributed to a number of factors including overfishing, loss of seagrass habitat, diminished water quality and microalgal blooms (Arnold 2001, Tettlebach &Wenczel 1993). In Florida, due to the decline in abundance of scallops, the commercial fishery has been closed entirely and the recreational fishery has been restricted. The loss of the commercial fishery has obvious economic implications, but the recreational fishing season has also been estimated by local non-governmental organizations to contribute at least $1 million per year to the local economy of the Steinhatchee and St. Joseph Bay area alone (B. Arnold, FWRI, personal communication). Furthermore, ecological impacts such as reduced water clarity may result from the absence of filterfeeding bivalves, such as the bay scallop, for which phytoplankton from the watercolumn is a main food source (Farias & Uriarte 2006).

Bay scallops live in sheltered inshore habitats (<12 m depth) with salinities above 20‰ and generally spend part of their life in seagrass beds. They recruit to the blades of seagrass as juveniles and inhabiting the seafloor as adults, where the seagrass is thought to provide some protection from predators (Thayer & Stuart 1974, Eckman 1987). Although the adults are not sessile like many other bivalve species, the distance over which they travel as adults is believed to be limited, thus dispersal and migration most likely occurs during the planktonic larval (veliger) stage (Knowlton & Jackson 1993). Generally individuals live no more than 18 months and will reproduce only once in a lifetime. The spawning period for *A. irradians* varies

throughout its range from June and July in the northern part of its range to late August through September in Florida (Barber & Blake 2006), during which an individual may release hundreds of thousands of gametes into the water column.

Bay scallops are simultaneous hermaphrodites which release both eggs and sperm during a single spawning event just minutes apart. Fertilization occurs externally, although some self-fertilization has been shown to occur in the reproductive tract (Wilbur 1995). Fertilized eggs develop into a planktonic larval stage that remains suspended for approximately two weeks, after which they will settle onto a substrate, preferentially on seagrass. Although as simultaneous hermaphrodites this organism seems well suited to low population numbers, densities below five scallops per 600 m^2 are thought to dramatically decrease recruitment of juveniles (Arnold *et al*. 1998). Because of this density-dependent limitation on recruitment, it may be difficult for depleted populations to recover following a catastrophic event, such as a red-tide (Peterson & Summerson 1992) or severe overfishing. On the other hand, as a species with a lifespan of 1-2 years, the population sizes tend to fluctuate naturally on an annual basis.

Despite restrictions on both commercial and recreational fishing implemented in 1995, Florida's depleted populations failed to recover. As a result, in 1999 restocking efforts were implemented using release of larvae from aquaculture broodstock with the intention of increasing the abundance of scallop populations in an area above the threshold necessary for spawning success. With aquaculture-based measures, a few individual scallops are collected from natural populations and spawned in an aquaculture facility. In some cases, the resulting larvae are grown to adulthood then deployed in cages into the wild to contribute to spawning. In other cases, the larvae are reared to a certain size and then released into containment booms, with the

expectation that they will settle, grow to adulthood, and spawn effectively in the wild by providing the necessary density of reproductively active individuals

Although a recent scallop management strategy, the stocking of some fisheries species has been a routine practice for over a century and has typically involved fish hatcheries releasing millions of fish into coastal waters (Liao *et al.* 2003). The application of similar strategies with a number of exploited invertebrate species have been attempted on a small scale with varying degrees of success (Burton & Tegner 2000). In any case, an important part of stock enhancement and restoration is the ability to quantify the impact on the species of interest. Enhancement of bay scallops in Niantic River estuary in 1998 revealed very little immediate increase in abundance (Goldberg *et al.* 2000) in contrast to the impressive increase in abundance of the same species in a transplant study conducted between 1992 and 1994 in Bogue Sound, North Carolina (Peterson *et al.* 1996), where adult density increased from less than $1/m^2$ to $15/m^2$. Such increases in the abundance of a species in a population following restoration may be attributed to those efforts, but direct evidence confirming the connection are often lacking (Wilbur *et al.* 2005). As some annual fluctuation in population size is common and expected in short-lived species like the bay scallop (Bologna 1998), the apparent resurgence in abundance following restoration may occur naturally. Consequently, it is not sufficient to document increases in abundance following restoration, the increases must be directly linked to the restoration effort. Determining the link between restoration and increase in abundance is possible through the use of genetic analyses.

Various studies of shellfish restoration have employed genetic tags for this purpose and have been successful in some instances. Using mitochondrial DNA, Milbury *et al.* (2004) confirmed the presence of offspring of restoration stock oysters in samples collected in

Chesapeake Bay. However, Wilbur *et al.* (2005) found that increased abundance of bay scallops at a number of restoration sites in Florida could not be directly attributed to the ongoing restoration efforts. Although genetic analyses are a useful tool for evaluation of restocking, they may also give conflicting or ambiguous results. For example, in an evaluation of red abalone (*Haliotis rufescens*) restoration in California, Gaffney *et al.* (1996) found a surprising deviation from Hardy-Weinberg equilibrium in a sample from the restoration site in 1992 and suggested that the deviation observed was due to success of the hatchery-produced restoration. Burton and Tegner (2000) sampled the same population in 1999 and did not find any significant deviation from expected allelic or genotypic frequencies, giving no evidence of success of the abalone restoration. Therefore, it is important to determine appropriate sampling strategy, genetic markers and genetic techniques for these evaluations. If the genetic marker is not diagnostic, sample sizes not sufficiently large, or genetic methods not powerful enough, they may be unable to provide convincing evidence as to the success of the restoration project. Evaluating the admixture of different stocks has been carried out on a much larger scale in fin fisheries, particularly with salmon, using multi-locus assignment methods, which have proven to be a powerful tool and may also provide an effective method for evaluating shellfish restoration.

CHAPTER 1: MICROSATELLITE MARKER DEVELOPMENT AND PARENTAGE ANALYSIS

INTRODUCTION

Microsatellites, also known as simple sequence repeats (SSRs) and short tandem repeats (STRs), are a relatively new class of molecular marker which have been used over the past decade for studies that use multi-locus and/or nuclear DNA analyses for assessing population structure. As genetic stock identification is the main goal of the data collected for this project, the molecular markers to be used would ideally be diploid, codominant and highly variable, all characteristics of microsatellites. However, there are other kinds of markers that could be used. Both allozyme analysis and mitochondrial DNA (mtDNA) sequencing work has been used to evaluate population structure and assess restoration in the bay scallop; however, each of these techniques has advantages and disadvantages that justify additional exploration into the utility of other markers.

Allozymes are enzyme molecules representing different variants of the same gene. Once they have been extracted from tissues they are separated by gel electrophoresis based on differences in size and electrical charge, as determined by their amino acid composition. This means that nucleotide-level variation resulting in no change of amino acid composition (such as many third-codon position substitutions) or in amino acid changes that do not affect the charge of the molecule will not be detected using this method. In addition, allozymes represent expressed regions of DNA and therefore also overlook significant variation found in unexpressed regions, such as introns (Kreitman 1983). Because allozymes are expressed proteins and therefore have some role in the life of an organism, it is likely that they are subject to selection.

As non-neutral markers their applicability to some studies is limited, but they are very useful for investigating how genetic variation affects the viability of individuals, populations and species.

Allozymes have been used to look at the effect of heterozygosity on somatic and reproductive growth in the bay scallop (Bricelj & Krause 1992), population structure (Bert *et al.* in prep), and stock identification (Krause 1992). Although allozymes have historically been used in stock identification in salmonid fisheries, they are slowly evolving markers with low variability and limited power to discriminate between recently diverged source populations (Smouse *et al*. 1994). In these cases, rapidly evolving markers with moderate to high polymorphism are more likely to provide the necessary power to discriminate between stocks. These types of markers include mtDNA RFLP and sequences, Single Nucleotide Polymorphisms (SNPs), Randomly Amplified Polymorphic DNA (RAPDs), and microsatellites.

Mitochondrial DNA is a single circular DNA molecule found in the mitochondria of eukaryotic cells. MtDNA is haploid, considered neutral, and generally maternally inherited (except in some bivalves e.g. *Mytilus edulis* (Hoeh *et al.* 2002)). Some genes in the mtDNA genome are highly conserved, and universal primers are available for commonly used mtDNA regions (Kocher *et al*. 1989). Mitochondrial DNA has become widely used because it displays a level of variation suitable for many different questions ranging from resolving interspecies phylogenetic relationships to intraspecies phylogeography and population structure. By obtaining nucleotide-level information through restriction fragment length polymorphisms (RFLP), denaturing gradient gel electrophoresis (DGGE), and especially sequencing, mtDNA is a sensitive marker for detecting unexpressed mutations, such as third codon position substitutions. While sequencing of nuclear genes is also possible, it is complicated by the diploidy of most animal nuclear DNA. The haploid state of mtDNA avoids the technical

complications of heterozygotes with two different sequences. Additionally, inefficient mutation repair mechanisms in mtDNA have been shown to contribute to the more rapid evolution of this genome (Brown *et al*. 1979), and studies have shown that in some cases mtDNA can provide higher resolution than even rapidly evolving nuclear DNA, such as microsatellites (Hoarau *et al.* 2004).

In bay scallops, mtDNA RFLP and sequencing has been used to look at both population structure (Blake & Graves 1995, Bologna *et al.* 2001, Bert *et al.* in prep) and the potential for genetic assessment (Wilbur *et al.* 2005). While mtDNA has been used for stock assessment in fisheries and population genetics (Epifanio *et al*. 1995, Bass *et al*. 2004), the limitations of this marker include uniparental inheritance and linkage of all mtDNA genes. As a result, mitochondrial markers do not reflect the contribution of half of the population (i.e. males) to a stock, and the lack of independence prevents the use of statistically powerful multi-locus analytical approaches. Two other kinds of markers may be applied to multi-locus analyses, RAPDs and SNPs. Some RAPD markers have been developed for *A. irradians* (Chikarmane *et al.* 2001); however, consistent scoring of these markers is difficult to achieve and most software designed to carry out multi-locus analyses does not support presence/absence data. While SNPs may be used for multi-locus analyses, it takes many more loci to obtain the same power of analysis as microsatellites because of the limited variability observed in SNPs (Hayes 2005).

Multi-locus analyses include assignment tests, mixed-stock analysis and parent-offspring analyses, in which likelihood-based methods are used to assign individuals to populations and offspring to parents (Manel *et al*. 2005). Microsatellites are excellent nuclear markers for multi-locus analyses because they are biparentally inherited, highly variable, tandemly repeated sequences (two to six basepairs long) found abundantly in eukaryotic genomes. Microsatellite

data is evaluated from isolating fragments that are known to contain variable number of repeated sequences, amplifying this same fragment for a number of individuals and comparing the banding pattern on an agarose or acrylamide gel or using an automated sequencer. Microsatellites have high mutation rates, generally on the order of 10^{-5}-10^{-2} mutations per locus per generation (Jarne & Lagoda 1996), though this rate may be dependent on the length of the repeated unit (di-, tri- or tetranucleotide), nucleotide composition of the repeat, allele length, and taxonomic group (Balloux & Lugon-Moulin 2002). These markers are thought to be neutrally evolving, as most are contained within non-coding DNA; however, some tri-nucleotide repeats have been found to be associated with diseases in humans and may also occur in exons where they do not disrupt the reading frame (Jarne & Lagoda 1996).

Due to the number and variability of microsatellite loci, they are a powerful tool for revealing subtle population structure (Shaw *et al.* 1999), determining kinship (Queller *et al.* 1993) and differentiating stocks (Beacham & Wood 1999, Beacham *et al.* 2004). One of the most important benefits of using microsatellite markers is the ability to use multiple independent loci simultaneously to assess individuals or populations. Increasing the number of loci used increases the statistical power of the assignments. Depending on the degree of polymorphism of the loci and the sample size of the study, it has been suggested that 7-9 microsatellite loci are appropriate to obtain high confidence of correct assignment in such studies (Zane *et al* 2002, Bravington & Ward 2004).

Consequently, microsatellites have been adopted as the marker of choice for evaluating the success of fisheries stocking programs, particularly for salmonid fishes (Kim *et al.* 2004, Ruzzante *et al.* 2004). The major obstacle to using microsatellites is that they are often species specific and therefore, unless markers have been developed for other studies, they must be

developed for the subject organism (Zane *et al*. 2002). Microsatellite markers previously had been developed for the sea scallop, *Placopecten magellanicus*, for parentage analysis and have shown high variability, indicating that this marker could be a powerful tool in this family of organisms (*Pectinidae*) (Gjetvaj *et al.* 1997). Concurrent to this study, Roberts *et al.* (2005) and Zhan *et al.* (2005) published microsatellite primers designed using expressed sequence tag (EST) databases. Some of the primers designed by Roberts *et al*. (2005) will be discussed in chapter 2 of this thesis; however, these loci appear to show less polymorphism than those designed in the present study.

In this study, the focus has been on developing tri- and tetra-nucleotide repeat loci. Loci with di-nucleotide repeats were not targeted for development because stutter peaks can cause ambiguity in scoring these loci, a problem much less common in tri- and tetra-nucleotide microsatellites.

MATERIALS AND METHODS

Development of Microsatellite Loci

Whole genome DNA was extracted from the adductor muscle of five individuals from three regions within the North American range of *A. irradians* (New York, North Carolina, and Florida) using a Puregene DNA Purification Kit (Gentra Systems, Minneapolis, MN). Genomic DNA was run out on 1% agarose gel stained with ethidium bromide to determine quality of the extraction. Ten microliters of each of the best four extractions from each population (three for Florida) were combined to make two DNA cocktails (two extracts per population each), which

were sent to Savannah River Ecology Laboratory, University of Georgia for microsatellite "double" enrichment. The procedure for this involves digesting high molecular weight DNA with restriction enzymes, cloning and amplification of the DNA, enrichment via hybridizing DNA fragments to specific biotin labeled SSR oligos, and elution of "enriched" DNA containing fragments with microsatellites (Kaukinen *et al*. 2004). Three oligo mixes (2, 3, and 4) containing different types and lengths of repeats were used to produce three different enrichments (Ai02, Ai03, and Ai04).

Each enrichment was PCR amplified using the following reagents: 1x Taq polymerase buffer, 2.0mM $MgCl_2$, 1 unit Taq polymerase (Promega, Madison, WI), 25µg/ml BSA, 200µM dNTPs, 0.5µM SNX-24f primer, 2µL eluted DNA enrichment, and sterile distilled water (dH_2O) to a total volume of 25µl. PCR was conducted using a PTC-200 thermocycler (MJ Research, Inc., Watertown, MA) and the following conditions: denature at 95^{o}C for 2 min., 35 cycles of 95^{o}C for 20 sec., 60^{o}C for 20 sec., 72^{o}C for 1.5 min., final extension at 72^{o}C for 30 min.. Ligation and transformation were conducted according to the pGEM-T Easy Vector System (Promega). Each transformation was plated on nine agar plates with ampicillin, using 50µl of culture per petri dish, and then incubated overnight (~16 hours). For each enrichment four hundred eighty positive colonies, indicated by white color, were picked using a sterile pipet tip, placed in 20µl dH_2O, boiled for 5 minutes, and stored at -80^{o}C.

Inserted fragments were isolated from the vector using PCR amplification using the following reaction: 1x Taq polymerase buffer, 1.5mM $MgCl_2$, 1 unit Taq polymerase (Promega), 150µM dNTPs, 0.4µM each SP6 and T7 primers, 0.5µL clone lysate, and dH_2O to a total volume of 25µl. Prior to amplification, clones were thawed to room temperature and centrifuged at 1000 rpm for 1 minute to pellet suspended cell remnants, and the supernatant was used for PCR

reactions. Amplifications were performed using both PTC-100 and PTC-200 thermocyclers (MJ Research, Inc.) and the following conditions: denature at 95°C for 5 min., 10 cycles of 95°C for 20 sec., 54°C for 30 sec., 72°C for 30 sec., 30 cycles of 95°C for 5 sec., 54°C for 30 sec., 72°C for 30 sec., final extension at 72°C for 5 min.. Each amplicon was sequenced in one direction using BigDye® Terminator (Applied Biosystems, Foster City, CA) and standard cycle sequencing conditions on PTC-100 or PTC-200 thermocyclers. Sequences were analyzed on an ABI3100 automated sequencer (Applied Biosystems), edited in Sequencher 4.1.4 (Gene Codes Corp., Ann Arbor, MI) and screened for presence and quality of tandem repeated nucleotide sequences by eye. Uninterrupted di-, tri-, and tetra-nucleotide repeats with consistent repeat units were sequenced in the reverse direction and aligned with sequences from other clones to make sure that loci were not repeated. Primers were designed to flank target regions as closely as possible using the Primer3 program (available online, http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi).

Primers were screened for consistent PCR amplification using samples from across a range of geographic populations (New York, North Carolina and Florida). Primer sets that amplified consistently were resynthesized with fluorescent label (either HEX or FAM, Applied Biosystems) on the forward primer and a 5' pigtail on the reverse primer to promote adenylation and minimize stutter peaks (Brownstein *et al.* 1996). Fluorescent-labeled primers were tested by PCR and optimized using various annealing temperatures and different concentrations of dNTPs and $MgCl_2$. After confirmation of amplification on an agarose gel, PCR products were diluted 1:10-1:20 with dH_20, and 1µl diluted PCR product was added to 9µl Hi-Di:Rox solution (with a ratio of 1025:25) and visualized on an ABI PRISM 3100 Genetic Analyzer. Resulting peaks were analyzed using GeneScan 3.7 and Genotyper 3.7 software (Applied Biosystems).

Evaluation of Microsatellite Loci

An issue in bivalves is the prevalence of heterozygote deficiencies in nuclear DNA markers. Among other potential causes, heterozygote deficiencies may be due to the presence of null alleles (alleles that do not amplify due to polymorphism in the flanking region of the locus and failure of primers to anneal to the template DNA). Although, null alleles have not been a consistent problem with bay scallops, as they have with other bivalves such as oysters, for example (Hedgecock *et al*. 2004, Hare *et al*. 1996), Mendelian inheritance of alleles was checked by genotyping a set of parents and offspring from an aquaculture broodstock spawn and performing parentage analysis.

Genotypes were collected for a set of six broodstock from Anclote Estuary in Florida (04psk1-6) and 56 of their offspring (04byop1-56) from a hatchery mass spawn (broodstock pooled during spawning). Parentage of the offspring was determined and a chi-square test was used to assess the goodness of fit of the genotypes of the offspring at each locus to Mendelian expectations given parental genotypes (Zar 1999).

Allelic richness, conformation to Hardy-Weinberg equilibrium (HWE) and genotypic (linkage) equilibrium in and among the five developed loci were examined in a total of 125 scallops collected from three locations in Florida (Anclote Estuary, Pine Island Sound and Steinhatchee). All data were scanned with MICRO-CHECKER 2.2.3 (Van Oosterhout *et al*. 2003) to identify any misrecorded alleles and to look for signs of null alleles, stutter peaks and large allele dropout which may have led to genotyping errors. Genetic Analysis in Excel 6 (GenAlEx) add-in (Peakall & Smouse 2005) was used to calculate expected and observed heterozygosity within populations. FSTAT 2.9.3 (Goudet 1995) was used both for analysis of

HWE and to estimate adjusted allelic richness. FSTAT adjusts allelic richness for sample size using a rarefaction method (El Mousadik & Petit 1996). Linkage disequilibrium was analyzed for each pair of loci within each population in GENEPOP 3.4 (Raymond & Rousset 1995). Results for HWE and linkage disequilibrium were assessed against a sequential adjusted Bonferroni *P*-value (Rice 1989).

RESULTS

The first enrichment to be cloned was Ai03 and was the source of the majority of the primers that have been produced by this effort to develop microsatellite markers. Of the 480 colonies picked from this cloning effort, all were amplified using PCR, 347 were sequenced at least in one direction, and of these, 34 did not work. Of the 313 clean sequences, 183 formed a contig with other sequenced fragments, while 152 were unique. There were a total of 61 potential microsatellite loci. This included 37 tetranucleotide (61%), 11 trinuclotide (18%), five dinucleotide (8%), and eight compound (13%) microsatellites, resulting in an overall 19.5% yield.

Enrichments Ai02 and Ai04 were also cloned, and fewer of these clones were amplified and sequenced because of the large number of clones already analyzed from the first enrichment. Of enrichment Ai02, 118 out of the 120 colonies were successfully amplified with PCR. Of the resulting 118 sequences, 46 formed a contig with other fragments, while 72 were unique. In total, there were 21 different microsatellite loci comprised of one tetranucleotide (0.5%), five trinucleotide (24%), and 15 dinucleotide (71.5%) microsatellites, resulting in a 17.5% yield. Of

enrichment Ai04, 48 colonies were amplified by PCR with two failures. Of the 46 sequences obtained from this enrichment, ten formed a contig with other fragments, while 36 were unique. A total of 11 microsatellites were observed, including three tetranucleotide (27%), three trinucleotide (27%) and five dinucleotide (46%) microsatellites, resulting in a 24% yield.

Of the 93 unique microsatellite sequences discovered, primers were designed for 77 loci. The remaining sequences either had repeat segments that were too long to be easily analyzed on the 3100 Analyzer (>500 bp), had highly irregular repeat units, or did not have sufficient flanking region for primers to be designed. Of those 77 loci for which primers have been designed, 34 sets were obtained and tested. A total of five primer pairs amplified consistently enough to be used for further studies (Table 1). Six of the 34 did not amplify at all, 11 sets amplified but inconsistently or produced too many bands or bands of the wrong size. Thirteen sets amplified well with unlabeled primers, but through analysis with labeled primers or sequencing, were either found to have a high frequency of null alleles, amplified nonmicrosatellite-containing products, or were monomorphic.

Genetic analysis using these five loci was sufficient to determine parentage for 48 of the 54 offspring samples (2 of the 56 offspring extracts did not amplify for any loci). Over 87% of those identified were the progeny of a single pair (scallops 04psk01 and 04psk05), and possibly an even larger portion, as the six individuals whose parents could not unambiguously be identified were either offspring of 04psk01 and 04psk05, or of one of these crossed with 04psk2 (Figure 1). Mendelian inheritance was confirmed for all of the five loci (P>0.05), although locus AICL327 was homozygous in both parents, and therefore completely homozygous in all offspring of this pair. In addition, although locus AICL327 showed heterozygosity in two of the six potential parents, all of the offspring were homozygous for the same allele (98 bp).

Evaluation of the allelic richness in samples of wild Anclote scallops and broodstock offspring revealed a dramatic loss of allelic diversity for all loci. Whereas in the 2001 wild Anclote Estuary population, for example, these five loci had 21, 12, 16, 9, and 6 alleles respectively out of 50 scallops (49 for AICL131), the 54 genotyped offspring (51 for AICL131) had 4, 4, 5, 5, and 1 alleles. In particular the low frequency alleles in the wild population were completely lost in the offspring, but additionally, dramatic shift in allele frequency of some moderate to high frequency alleles has occurred (Appendix A).

Comparison of allelic richness shows comparable values among all three wild populations (Table 2). Hardy-Weinberg equilibrium was met for most loci at all three populations; however, AICL112 and AICL115 showed evidence of heterozygote deficiency in a single population each (Table 2). No linkage disequilibrium was observed among loci within any of the three populations or when populations were pooled.

Table 1. Summary information for microsatellite loci. Bracketed basepairs identify pig-tail (Brownstein *et al.* 1996) added to enhance resolution of genotypes.

Locus	Primers	Size (bp)	Annealing Temp. (^{o}C)	Repeat
AICL112	F: TGCCAAATCCATTTGCATATTA R: [GT]TTCCCTGTTCACTTGACAGACC	214	56	$(GACA)_1GATG(GACA)_{12}$
AICL115	F: TGCGGTATTTGAGTCCCCTA R: [GT]TTGACCTTTTGACCCCAAAT	201	56	$(GTCT)_{10}$
AICL131	F: CCCTATGGCTTCCTCAACCT R: [GT]TTAACTTTCTGTGCCGTGGA	250	50	$(CAA)_9$
AICL271	F: CCTTACATGACCCTGGCTGT R: [GT]TTCATCTAATTTATCAACCGACCA	91	50	$(CAAA)_8$
AICL327	F:GCAAAATCCACCCATCAGTT R:[GTTT]ACCGGAGGGGACTAGTGTTT	103	58	$(CAGA)_6$

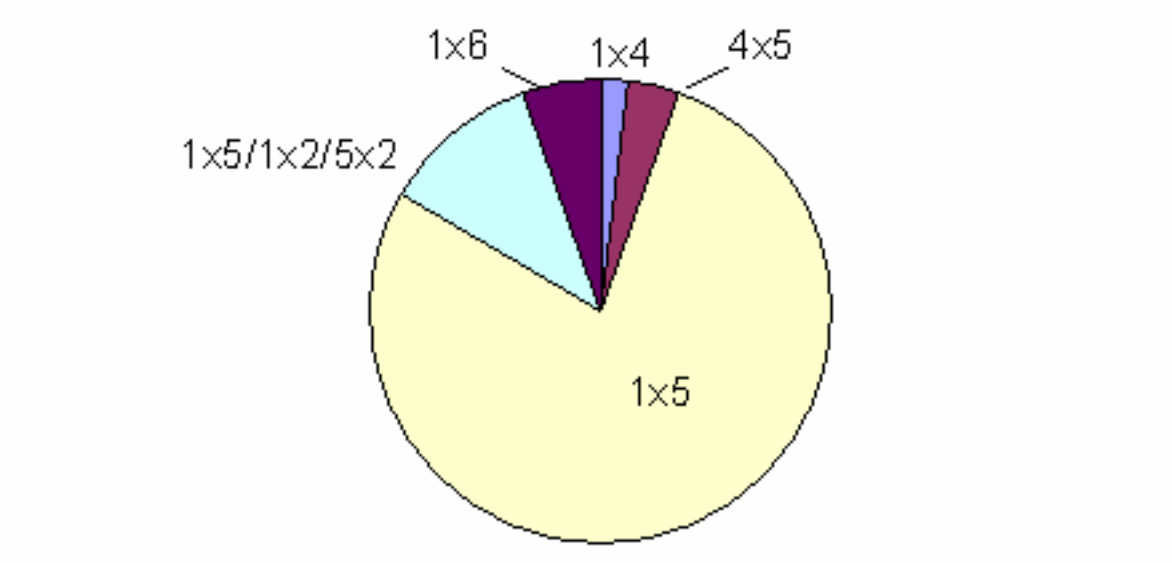


Figure 1. Parentage of offspring as determined by microsatellite analysis.

Table 2. Statistics of microsatellite loci: number of individuals scored (N), number of alleles observed (N_a), allelic richness (R) observed heterozygosity (H_o), expected heterozygosity (H_e), and *P*-values for heterozygote deficiency (* indicates significance at adjusted *P*-value) for Anclote Estuary (AN), Pine Island Sound (PI) and Steinhatchee (ST) samples.

Locus	Sample	N	N_a	R	H_o	H_e	*P*
AICL112	2001AN	50	21	14.366	0.760	0.863	0.0120*
	2001PI	50	12	9.981	0.800	0.831	0.2435
	1998ST	25	13	13.000	0.800	0.844	0.2435
AICL115	2001AN	50	12	8.972	0.600	0.707	0.0204
	2001PI	49	10	8.655	0.612	0.652	0.2000
	1998ST	25	7	7.000	0.400	0.614	0.0028*
AICL131	2001AN	49	16	12.598	0.614	0.818	0.0102
	2001PI	50	13	11.445	0.840	0.862	0.3037
	1998ST	25	11	11.000	0.920	0.825	0.9417
AICL271	2001AN	50	9	8.445	0.840	0.788	0.8574
	2001PI	50	9	7.859	0.800	0.760	0.7796
	1998ST	25	7	7.000	0.800	0.714	0.8787
AICL327	2001AN	50	6	4.876	0.300	0.303	0.5333
	2001PI	50	6	4.878	0.280	0.271	0.7157
	1998ST	25	5	5.000	0.240	0.222	1.000

DISCUSSION

Microsatellites have become a popular tool for studies of population genetics of many organisms, and development of these markers is a necessary step prior to their application for a range of purposes. It is a laborious, expensive, and time intensive endeavor, however, and in this case produced a yield of about 20% of clones possessing a novel microsatellite, and 83% of those were able to have primers designed. Of the 34 loci tested, five (14.7%) resulted in viable loci that will be applied in the second part of this study. Although a yield of 14.7% seems quite low, it is nonetheless consistent with other attempts to design microsatellites for this species. For example Zhan *et al.* (2005) found 11 polymorphic and applicable microsatellites after screening 66 sequences with repeat motifs. This is a yield of 16.7%. Roberts *et al.* (2005) had a somewhat higher yield of eight polymorphic loci from 29 SSR containing sequences (27.6%). As the EST derived microsatellites seem to have lower variability than those developed through use of an enriched microsatellite library, it may be that their flanking regions, and thus primer sites, are more highly conserved as well. Whereas microsatellites recovered from an enriched library may originate anywhere in the genome, those recovered from an EST database are found in expressed genes or their introns, possibly constraining their variability. Thus, there may be a trade off between more loci and more variable loci when using different isolation techniques.

It should be kept in mind, that different taxa have very different yields of microsatellites and that even within taxa the degree of success is highly variable (Zane 2002). Generally fish and some other vertebrates have higher yields of microsatellites than invertebrates and plants. In particular, when specifically searching for tri- and tetra-nucleotide microsatellites, yields can be especially low due to their relative infrequency in the genome of many organisms (Kaukinen *et al*. 2004). A consequence in this study for not targeting dinucleotide repeats for development

may be a lower success rate than had all types of repeats been equally considered for development.

Following the discovery of potential microsatellites there are a number of other criteria to consider when determining their applicability to a project. In particular, null alleles were a significant problem in the development process, as a number of potential microsatellite loci amplified inconsistently, or showed evidence of null alleles. Using a set of parents and offspring to look at parentage and Mendelian inheritance was found to be a useful tool for identifying unreliable loci prior to genotyping full populations of individuals. Null alleles can cause problems when conducting analyses that assume Hardy-Weinberg equilibrium because they can contribute to heterozygote deficits. Heterozygote deficits may also be a result of Wahlund effect, or the mixing of two genetically distinct populations. Because of these two different sources of Hardy-Weinberg disequilibrium, it is wise to eliminate loci that can be shown to have null alleles by non-Mendelian inheritance. The markers presented here do not show ubiquitous heterozygote deficiencies or excess in the populations surveyed thus far, only loci AICL112 and AICL115 show heterozygote deficiency in a single population each.

Aside from providing a method for assessing Mendelian inheritance of the loci, parentage analysis also revealed that multi-locus analysis can provide a higher degree of resolution in determining parentage than analysis with a single mitochondrial marker. Based on mitochondrial sequences, evaluation of parentage indicated that 31 offspring (70.5%) were the result of a single mtDNA donor (04psk01), while the remaining 13 were split among three other mtDNA sources (three from 04psk04, three from 04psk06 and seven from 04psk05). The other two potential parents were not represented by mtDNA in the offspring sample. Microsatellite analysis consistently supported the mtDNA evidence of parentage and furthermore was able to

identify both parents in 89% of the offspring. The microsatellites revealed that in a spawn of a small number of individuals, only one pair of scallops sired at least 87% of the identified offspring (42 individuals), and only four of the six potential parents contributed at all. It is somewhat surprising that despite the small number of scallops used in the spawn there is no evidence of self-fertilization. Bay scallops, like many pectinids are simultaneous hermaphrodites and it may be expected that some self-fertilization would occur in hatchery situations where there are a limited number of spawning individuals, or even when there is a low density of scallops in the wild.

Some of the loci were much more useful for determining parentage than others. AICL327, which is the least polymorphic of the five loci, was represented by only two alleles in the parents and was completely monomorphic in the offspring, despite the contribution of four of the six parents to the gene pool. This demonstrates that some loci (particularly those with few alleles) may not be helpful in parentage analysis. This is consistent with the findings of Bernatchez and Duchesne (2000), which indicated that the allocation success is dependent on both number of loci and the variability of the loci, and that loci with high allelic diversity are best for parentage analysis and those with moderate allelic diversity (number of alleles between six and ten) are best for population assignment. Furthermore, spawning of small numbers of parents with few alleles at a locus can severely reduce or eliminate genetic diversity at that locus in subsequent generations. Causes of the dramatic genetic drift in a single generation could be failure of individuals to spawn in synch with the other adults or failure to spawn at all, pre-zygotic selection against sperm and eggs of certain individuals, and post-zygotic selection against unfit offspring. This dramatic shift in gene pools poses both a potential opportunity and a serious concern for restoration projects.

While opportunity may arise from the potential to use this shift in allele frequency as an identifier for stocked scallops and thus facilitate the assessment of success, concern arises from the loss of genetic diversity in a population subsequent to aquaculture-based restocking efforts. Loss of genetic diversity, particularly in small populations, can lead to reduced adaptability, survival and reproduction of the population overall (Frankham *et al.* 2002). Concerns also arise over using restoration organisms from disconnected populations because of the threat of introducing genetically dissimilar organisms and shifting the genetic diversity of the native population. As this study indicates, a broodstock originating in the same population can have a dramatically different genetic profile than the wild population, and may be as dissimilar to the wild as individuals from another population. This is not particularly surprising, as genetic drift and loss of low frequency alleles has been demonstrated in a number of aquaculture systems including abalone (Evans *et al.* 2004), Pacific oyster (Boudry *et al.* 2002, Hedgecock & Sly 1990), flat oyster (Launey *et al.* 2001) and European oyster (Saavedra 1997). This effect should be taken into consideration when planning restoration projects, as spawning batches are generally comprised of six to eight spawning individuals (Arnold 2001) and may have similarly biased contribution to reproduction.

CHAPTER 2: ASSIGNMENT AND CLUSTERING OF BAY SCALLOPS IN FLORIDA AND THE ATLANTIC

INTRODUCTION

Pine Island Sound is an estuarine system on the southwest coast of Florida with freshwater input from the Caloosahatchee River. Up until the late 1980s bay scallops supported a recreational fishery in the area, but subsequently numbers have declined leaving only a small population in the northern end of Pine Island Sound, a shallow area with abundant turtle grass (*Thalassia testudinum*). In surveys conducted between 1995 and 2002, the Florida Fish and Wildlife Research Institute found very few individuals remaining in Pine Island Sound, with density consistently below six scallops per $600m^2$ (B. Arnold, FWRI, personal communication). In 2003, one and a half million competent larvae produced in a hatchery using broodstock from Anclote Estuary were released into three construction boom enclosures. Spat were found to have settled only in the enclosures that received restoration larvae. In 2004, transect surveys observed a density of adult scallops in the restoration area that were two orders of magnitude greater than densities found in the other areas of Pine Island, to 136-192 scallops/$600m^2$ (Leverone 2004). In 2005, surveys showed even greater abundance of scallops in the sound (B. Arnold, FWRI, personal communication). However, the increase in abundance has not yet been genetically tied to the larval release.

As molecular data has become more readily available, new statistical methods have been developed and are being applied to understand migration, genetic introgression and population structure in fisheries as well as a wide range of organisms. Statistical approaches that are well suited to the assessment of admixed populations are assignment methods and Bayesian cluster

analysis. Assignment methods consist of mixed-stock analysis and assignment tests and use allele frequencies of populations determined *a priori* by the researcher as baseline data.

Mixed-stock analysis (MSA) is a statistical method originally employed in fisheries using physiological markers including otolith and parasite analyses as well as meristic and morphometric character analysis to determine the contribution of origin populations to a mixed population (Reynolds & Templin 2004). More recently, MSA has become synonymous with genetic stock identification (GSI), a term applied to statistical methods using molecular genetic data to determine the contribution of source populations to the mixed stock using allele frequencies of the presumed source populations. Molecular MSA can be used to understand the magnitude of migration between geographic areas or introgression of genes into a population due to stock enhancement and is especially valuable as it can be applied in cases where there is little physiological difference between the stocks. Molecular MSA can use either mitochondrial or nuclear markers and has been extensively used to study populations of a number of fisheries species.

However, whereas MSA attempts to determine the overall contribution of different stocks to a single sample, assignment tests consider the multi-locus genotype of individuals and assign or exclude them from sources using probability to determine which population is the more-likely source of the individual (or which populations are most likely not the source) (Cornuet *et al.* 1999). Because assignment tests address origin at the scale of the individual rather than the population, they calculate likelihood values for the origin of every individual and leave confidence of the assignment to the discretion of the researcher. Two main statistical methods have been developed to calculate these likelihoods: a frequency-based method originally applied to evaluate population structure in polar bears (Paetkau *et al*. 1995) and a Bayesian method

initially applied to investigating recent migration between human populations (Rannala & Mountain 1997). Paetkau *et al.* (1995) introduced a frequency-based method wherein the likelihood of an individual's multi-locus genotype originating in each of the potential source populations is the product of the likelihoods of the genotypes of each locus in each population and based on the Hardy-Weinberg principle (probability of a homozygous genotype is p^2 and heterozygous genotype is 2pq, where p and q are the allele frequencies in the sample). The Bayesian method of Rannala and Mountain (1997) is similar, but derives a probability distribution of genotypes in the populations using a Bayesian algorithm.

Following the widespread use of these methods to study various systems, Cornuet *et al.* (1999) introduced a genetic-distance-based method. This turned out to be less powerful than the other two methods; however, Cornuet *et al.* (1999) also described a probability-based method for excluding populations as sources, which has become applied in conjunction with both the frequency and Bayesian assignment tests. The theoretical difference between this last method and the traditional assignment test is the underlying question it addresses. Whereas traditional assignment methods simply answer, "which population is the more likely source of this individual?", the exclusion method addresses the possibility of unsampled source populations and answers the question "how probable is it that this individual originated in population X?".

Over the past ten years, assignment tests have been used much the same as MSA, to examine source populations of fisheries species such as Atlantic salmon, to understand population dynamics (Martinez *et al.* 2001, Vasemagi *et al.* 2001), and to detect fishing competition fraud or illegal poaching (Primmer *et al*. 2000). Assignment tests have also continued to be applied outside of fisheries to look at population structure and migration in many species. Recently Hare *et al.* (2005) used assignment tests to assess the success of eastern oyster

(*Crassostrea virginica*) restoration in the Chesapeake Bay. In this case, an artificially selected disease resistant strain of oysters (DEBY) was used to supplement the wild population. The following year, using eight microsatellite loci and a mitochondrial DNA marker, a number of wild-caught juveniles were identified as F_1 offspring of DEBYxWild oysters. This was achievable due to the strong effect of genetic drift at neutral microsatellite loci in the selected DEBY strain compared to the native Chesapeake Bay oysters, resulting in easily distinguishable genotypes of restoration oysters. In the case of clearly differentiated stocks, assignment methods appear to be a promising tool for assessing bivalve restoration, much as they have been historically used to reveal effects of restocking of finfisheries.

The effectiveness of assignment methods to distinguish between native and introduced stocks and their progeny depends on three critical factors: (1) inclusion of all potential source populations in the analysis, (2) adequately informative and numerous molecular markers, and (3) sufficient genetic homogeneity within and differences among stocks (Hansen *et al.* 2001). Cornuet *et al.* (1999) suggest that as an estimator of population differentiation F_{ST} can be a useful predictor of the performance of assignment methods. It is expected that assignment tests perform well when $F_{ST} \geq 0.05$, and can provide 100% accuracy when $F_{ST} \approx 0.1$, given use of ten loci, and a sample size of 30-50 individuals from each of ten populations. However, more loci (>20) and larger sample sizes (50 individuals) are necessary to achieve maximum accuracy with exclusion methods given ten populations and the same degree of population differentiation.

The second statistical method for assessing admixed stocks, Bayesian cluster analysis, is not technically an assignment method because it does not consider *a priori* determined populations. Bayesian clustering groups individuals so as to minimize linkage disequilibrium and Hardy-Weinberg disequilibrium within each cluster. Linkage disequilibrium and Hardy-

Weinberg disequilibrium occur when populations with different allele frequencies are combined in analysis. Like the aforementioned assignment tests, the power of Bayesian clustering for detecting structure depends on the magnitude and distribution of genetic variation within the samples. Due to the homogenization of allelic frequencies when there is high gene flow and low differentiation between clusters (F_{ST}<0.05), Bayesian clustering will have low power in detecting genetic structure.

Bayesian clustering has become a popular tool in studies that also use assignment methods because of the uncertainty in designating which groups of samples are discrete populations (Fraser & Bernatchez 2005, Eldridge *et al.* 2001). Furthermore, Berry *et al.* (2004) used Bayesian clustering to look at its ability to measure dispersal compared with Bayesian assignment tests and found the two methods to be comparably powerful.

This study evaluates the utility of these statistical approaches in conjunction with microsatellite analysis for the assessment of scallop restoration. This evaluation was conducted in two parts. First, microsatellite data (using the five loci described in chapter one and four loci described in Roberts *et al*. (2005)) was collected for three samples obtained from New York, North Carolina and Florida (Figure 2). Assignment tests were applied, based on these nine loci, to determine the ability to accurately identify the origin of individual scallops. Bayesian clustering was applied to further assess the genetic structure of the populations. The second part of the evaluation focused on the utility of these approaches for assessing the impact of bay scallop restorations, specifically the effect of the 2003 release of larvae derived from Anclote Estuary scallops into Pine Island Sound (Figure 3).

Figure 2. Map of regional scallop collection sites with range of *A. irradians* subspecies, *A.i. irradians* (light grey), *A.i. concentricus* (dark grey), and *A.i. amplicostatus* (black).

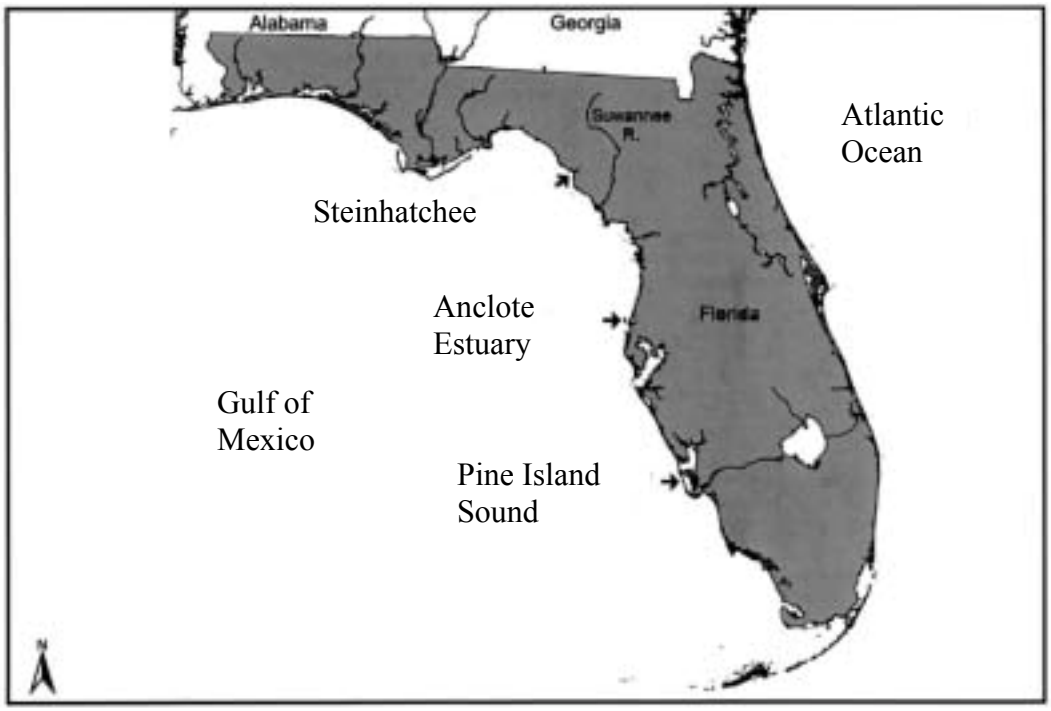


Figure 3. Map of Florida scallop collection sites.

MATERIALS AND METHODS

Data Collection

Nine microsatellite loci were used for the following analyses, including the five loci developed and described in chapter one in addition to four loci developed by Roberts *et al.* (2005, Table 3). Bay scallop samples from three populations separated by large geographic distances, and presumed to be genetically dissimilar were genotyped to assess the utility of these microsatellites in assignment tests and Bayesian cluster analysis. Samples contained approximately 50 scallops each originating in Florida (collected in 2001 in Anclote Estuary), North Carolina (collected in 1998 from Bogue Sound and Core Sound); and New York (collected in 1999).

For the evaluation of the restoration, the potential sources for the post-release sample of 50 scallops (2005PI) were either wild local scallops or the hatchery-produced larvae released in fall 2003. The actual broodstock used to generate the larvae that were released were unavailable for analysis, and thus a sample of 50 scallops from Anclote Estuary (2001AN) were used to approximate the genetic composition of the larvae. The wild Pine Island Sound source was characterized using a sample of 50 scallops from Pine Island Sound prior to the restoration in 2001 (2001PI). Twenty-five individuals from Steinhatchee, Florida, collected in 1998 (1998ST), were genotyped to account for another potential source population. Finally, 50 scallops collected in Pine Island Sound post-restoration in 2005 (2005PI), were genotyped for the purpose of determining if they could be assigned to any of the potential source populations.

Extraction and amplification of samples was conducted as described in chapter one. Again, fluorescently labeled PCR products were diluted (1:10 – 1:100) according to the intensity of the product when electrophoresed through a 2% agarose gel stained with ethidium bromide. One microliter of diluted PCR product was run through an ABI PRISM 3100 Automated Sequencer along with ROX dye standard, and resulting peaks were analyzed using GeneScan 3.7 and Genotyper 3.7 software (Applied Biosystems). Non-specific peaks, identified as those occurring regularly as a third peak, were not scored in data. Two loci required binning of alleles prior to further analyses: locus AICL112 (a tetranuclotide repeat) revealed fragments differing by two basepairs rather than the expected four basepairs, suggesting the presence of an insertion/deletion mutation somewhere in the fragment, and locus N391 (a dinucleotide repeat), which was difficult to score due to variation in peak size. For both loci, adjacent size classes were combined (binned) to generate allele classes that differed by four basepairs (i.e. all peaks 223-226.99 bp were called allele 225). Allele distributions for all samples were evaluated using MICRO-CHECKER 2.2.3 (Van Oosterhout *et al.* 2003) to identify any misrecorded alleles and to look for signs of null alleles, stutter peaks and large allelic dropout which may have led to genotyping errors.

Table 3. Microsatellite loci developed by Roberts *et al.* (2005). Bracketed basepairs identify pig-tail (Brownstein *et al.* 1996) added to enhance resolution of genotypes.

Locus	Primers	Size (bp)	Annealing Temp. (^{o}C)	Repeat
M26	F: CACTTTCAGCAGATATTCTTGAGG R: [GTT]TCCCATCCTCTCCTTCACAG	123	55	$(GAT)_{10}$
G340	F: CGCTTGTGTTTTACGAGGAGAAGG R: [GTT]TGACGGGGTGTGATGTCTGACC	117	53	$(GAT)_5$
S336	F: GCGGAGGCAGATTCTTTCTTTTC R: [GTTT]GGTCGTGGATTGTAAGCATTGTC	132	54	$(CAG)_5$
N391	F: TCATCGCCTCCACCTTCAG R: [GTTT]GATCACACTTTGATTTGTCCTACG	247	58	$(AG)_{14}A(AG)_5$

Data Analysis

All data were analyzed using the Genetic Analysis in Excel 6 (GenAlEx) add-in (Peakall & Smouse 2005) to determine allelic frequencies in each population. GENEPOP 3.4 (Raymond & Rousset 1995) was used to calculate deviations from linkage equilibrium. Hardy-Weinberg disequilibrium, number of alleles and allelic richness adjusted for sample size (El Mousadik & Petit 1996) were calculated in FSTAT 2.9.3 (Goudet 1995). Hardy-Weinberg disequilibrium and linkage disequilibrium were assessed against an alpha level of 0.05 adjusted with the sequential Bonferroni procedure (Rice 1989).

Estimators of population differentiation, F_{ST} (θ, Weir & Cockerham 1984), and R_{ST} (ρ, Rousset 1996 as estimated by Michalakis & Excoffier 1996), were calculated in GENEPOP. Both F_{ST} and R_{ST} estimators were calculated because they determine population differentiation based on two different mutation scenarios of microsatellites. F_{ST} calculates genetic differentiation assuming the Infinite Alleles Model (Kimura & Crow 1964), in which any allele can mutate directly to any other allele; whereas R_{ST} is based on the Stepwise Mutation Model (Kimura & Ohta 1978) in which an allele can mutate only by adding or deleting a single repeat unit. Due to the high variance associated with R_{ST}, F_{ST} is expected to give more accurate results when sample size and number of loci are small. However, studies have shown R_{ST} to be a better estimate of differentiation when separation between samples is large (Balloux & Lugon-Moulin 2002). This is because, as populations diverge over time, the effect of mutation becomes more important than migration in determining the extent of differentiation. GENEPOP was also used to perform nonparametric exact tests of genic population differentiation (Markov chain method), which are expected to be highly sensitive to low levels of divergence.

To determine the potential power of assignment tests, GenAlEx was used to conduct assignments using the frequentist method (Paetkau *et al.* 1995). Individuals were assigned to population of origin in using the leave-one-out procedure (Waser & Stroebeck 1998), which eliminates the bias of assigning individuals to populations of which they have contributed to baseline data. A default allele frequency of 0.01 was applied for alleles not present in a sample. Accuracy of assignment tests was assessed by the determining the number of individuals that were correctly assigned to their population of origin. In addition, the genotype likelihood ratio distance, D_{LR} (Paetkau *et al*. 1997), was calculated as the average value of the log-likelihood differences within a sample. Where $D_{LR} \geq 1$, 2 or 3, the genotypes of individuals from the two populations being compared are, on average, 10, 100 or 1000 times more likely to occur in their true source population than the alternate source.

The software package WHICHLOCI 1.0 (Banks *et al.* 2003) was used to rank all nine loci for their power of assignment. This program generates randomly resampled populations of equal sample size (N=100) and similar allelic frequencies to the original baseline populations and then performs single-locus assignments. In addition, it performs assignments using combinations of loci to determine the number of loci necessary to achieve a threshold accuracy (number of correct assignments) and stringency (relative likelihood) of assignment.

Bayesian clustering of individuals was conducted using STRUCTURE 2.0 (Pritchard *et al.* 2000), a burnin period of 50,000 with a run time of 100,000, an admixture model and putative population information included. The program was run for a range of groups (K) from one to five.

For the assessment of the 2003 larval release, assignment of 2005PI scallops was performed using GENECLASS2 (Piry *et al.* 2004) to apply frequency-based (Paetkau *et al.*

1995) and Bayesian (Rannala & Mountain 1997) assignment tests. Individuals were assigned to the population with the highest (least negative) log likelihood. Three levels of stringency were applied to these assignments where relative likelihood of assignment to the first ranked population (A) compared to the second ranked population (B) Λ=[-log10L(popA)]-[-log10L(popB)] is 1, 2, or 3, representing 10, 100, and 1000 times greater likelihood of originating in population A than B (when Λ=0 the likelihood of originating in either population is equal; Hare *et al*. 2005). In addition, probability-based exclusion tests were performed using GENECLASS2 with probabilities computed using a Monte-Carlo resampling procedure, the simulation algorithm of Cornuet *et al*. (1999) and 1000 simulated individuals. Exclusion tests were computed using three levels of stringency: $P \leq 0.05$, $P \leq 0.01$, and $P \leq 0.001$.

RESULTS

Applicability of Assignment Tests and Bayesian Clustering

No significant differentiation of North Carolina populations (Bogue Sound and Core Sound, F_{ST}=0.0065), permitted the pooling of these samples as a single region (Potvin & Bernatchez 2001) for the purpose of comparison with samples from the two other regions, Florida (represented by the 2001 Anclote sample) and New York. The Florida sample was represented by much higher allelic diversity than the other populations for three loci (AICL112, AICL115 and AICL131), even after correction for sample size. A significant heterozygote deficiency was observed for one locus (AICL112) in the Florida population (Table 4).

Genotypic disequilibrium was significant for one pair of loci in Florida (AICL112/G340, adjusted $P<0.0167$).

F_{ST} and R_{ST} estimators of population differentiation indicated significant differences between each of these three regions. However, there was a discrepancy between these two measures of differentiation. Whereas the largest difference for F_{ST} was observed between Florida and North Carolina and the least difference was between North Carolina and New York, R_{ST} showed the least difference between Florida and New York (Table 5). F_{ST} calculated by pairwise comparisons of the regions per locus indicated that the greatest structure was observed for locus AICL327 (0.1629-0.6148), followed by S336 (0.353-0.1465), N391 (-0.0061-0.898) and G340 (0.0223-0.0801). Genic differentiation was highly significant for all loci in Florida/North Carolina comparison, all loci except AICL115 in Florida/New York comparison, and was significant for AICL115, AICL327, M26 and S336 in North Carolina/New York comparison (Table 6, allelic frequencies in Appendix B).

Assignment tests were moderately accurate with samples divided into three populations (Florida, North Carolina, and New York). All Florida scallops were correctly assigned, but 27% of North Carolina scallops were misassigned to New York and 21% of New York scallops were misassigned to North Carolina (Figure 4, Table 7). No scallops were assigned from either the New York or North Carolina population to the Florida population. The average log-likelihood difference (D_{LR}) for population assignment of the Florida and North Carolina samples and Florida and New York samples were 4.154 and 4.639, indicating the likelihood of the genotype of an individual to have originated in their own population was more than four orders of magnitude greater than the likelihood of originating in the other population. In assignment of North Carolina versus New York samples, D_{LR} was 0.694.

Using all loci, WHICHLOCI analysis was unable to achieve 95% accuracy of assignment (at stringency Λ=0) when New York and North Carolina populations were considered separately, yet based on all loci was able to achieve 95% assignment accuracy at a stringency Λ=2 when North Carolina and New York populations were pooled. Furthermore, the same level of accuracy was achieved at a stringency of Λ=0 using a single locus (AICL327), and at a stringency of Λ=1 using two loci (AICL327 & N391). The most diagnostic locus on a regional scale by far was AICL327, which was consistently ranked first in WHICHLOCI analysis. When this one locus was removed from the analysis, assignment accuracy dropped off dramatically and became highly variable. To achieve a 95% accuracy of assignment, stringency dropped to Λ=0, at this level, some iterations of the assignment test found between four and seven loci sufficient, while other iterations were unable to achieve this accuracy using all eight loci. AICL112 and N391 were generally ranked either second or third and the other six loci ranked at various positions for each round of analysis.

Bayesian cluster analysis in STRUCTURE also indicated two likely populations (Table 8); although the log likelihood value was marginally smaller for three groups, Pritchard and Wen (2004) recommend choosing "the smallest value of K that captures the majority of the data". Two groups is consistent with the visual representation, which shows two separate clusters: one for individuals originating in New York and North Carolina and a second for those originating in Florida (Figure 5).

Table 4. Statistics of microsatellite loci in regions. Number of individuals scored (N), number of alleles observed (N_a), allelic richness (R) observed heterozygosity (H_o), expected heterozygosity (H_e), and *P*-values for heterozygote deficiency in regional samples. (*indicates significance at adjusted *P*-value)

Locus	Sample	N	N_a	R	H_o	H_e	*P*
AICL112	FL	50	21	20.274	0.760	0.863	0.0120*
	NC	49	16	15.834	0.939	0.897	0.8444
	NY	47	14	14.000	0.809	0.882	0.0778
AICL115	FL	50	12	11.694	0.600	0.707	0.0204
	NC	48	8	7.958	0.708	0.699	0.5667
	NY	47	9	9.000	0.681	0.675	0.5648
AICL131	FL	49	16	15.792	0.614	0.818	0.0468
	NC	49	10	9.913	0.653	0.716	0.1574
	NY	47	8	8.000	0.681	0.687	0.4889
AICL271	FL	50	9	8.997	0.840	0.788	0.8574
	NC	49	8	7.917	0.735	0.696	0.8037
	NY	47	6	6.000	0.638	0.660	0.3556
AICL327	FL	50	6	5.880	0.300	0.303	0.5333
	NC	49	4	3.918	0.347	0.419	0.1222
	NY	47	6	6.000	0.660	0.625	0.7056
	FL	50	5	4.940	0.800	0.666	0.9898
M26	NC	49	5	4.959	0.592	0.644	0.1907
	NY	47	6	6.000	0.809	0.712	0.9519
	FL	50	7	6.880	0.480	0.550	0.0630
G340	NC	49	6	5.959	0.816	0.714	0.9611
	NY	47	5	5.000	0.660	0.567	0.9648
	FL	48	3	3.000	0.458	0.520	0.2056
S336	NC	48	4	3.979	0.479	0.512	0.3111
	NY	47	3	3.000	0.681	0.630	0.7870
	FL	50	10	9.997	0.860	0.859	0.4778
N391	NC	49	10	9.877	0.755	0.712	0.8148
	NY	47	8	8.000	0.681	0.672	0.5463

Table 5. F_{ST} (θ) and R_{ST} (ρ) coefficients between regions Florida (FL), North Carolina (NC) and New York (NY) for each locus and for all loci.

	FL/NC		FL/NY		NC/NY	
Locus	F_{ST}	R_{ST}	F_{ST}	R_{ST}	F_{ST}	R_{ST}
AICL112	0.0250	0.2211	0.0324	0.1112	-0.0025	0.0338
AICL115	0.0139	-0.0102	-0.0065	-0.0092	0.0146	-0.0121
AICL131	0.0147	-0.0091	0.0083	0.0191	0.0177	0.0123
AICL271	0.0113	0.0001	0.0314	-0.0023	-0.0007	-0.0130
AICL327	0.6148	0.4277	0.5144	0.0020	0.1692	0.2554
M26	0.0136	0.0389	0.0681	0.0105	0.0813	0.0878
G340	0.0731	0.0710	0.0194	-0.0025	0.0296	0.0319
N391	0.1214	0.0743	0.1409	0.0701	-0.0061	-0.0089
S336	0.0673	0.0031	0.848	0.0466	0.0353	0.0855
All	0.1211	0.1411	0.1137	0.0545	0.0354	0.0567

Table 6. Pairwise regional genic differentiation probabilities per locus with Standard Error (*indicates significance at $P<0.05$).

Locus	FL/NC	FL/NY	NC/NY
AICL112	0*	0*	0.551 ± 0.014
AICL115	0.022 ± 0.004*	0.569 ± 0.012	0.044 ± 0.006*
AICL131	0.007 ± 0.001*	0.014 ± 0.002*	0.089 ± 0.008
AICL271	0.010 ± 0.002*	0*	0.210 ± 0.008
AICL327	0*	0*	0*
M26	0.084 ± 0.005	0*	0*
G340	0*	0.010 ± 0.002*	0.073 ± 0.005
N391	0*	0*	0.825 ± 0.007
S336	0.001 ± 0.001*	0*	0.005 ± 0.001*
All	*	*	*

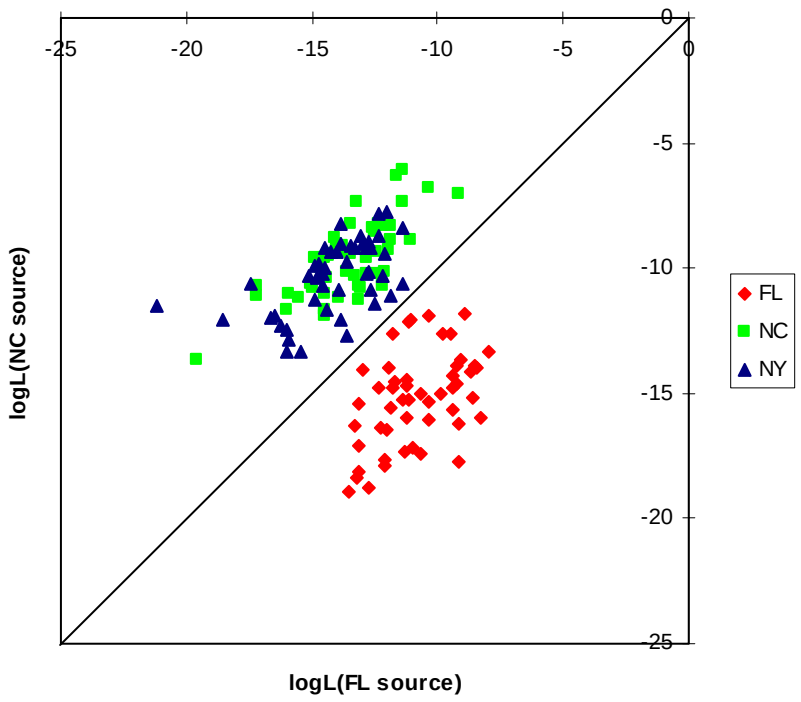


Figure 4. Assignment results for three regional populations. Points to the right of the line indicate assignment to FL, those to the left are assigned to NC.

Table 7. Regional assignment of samples to known source populations.

Source	Assigned to			% Correct
	FL	NC	NY	
FL	50	0	0	100
NC	0	36	13	73.5
NY	0	10	37	78.7

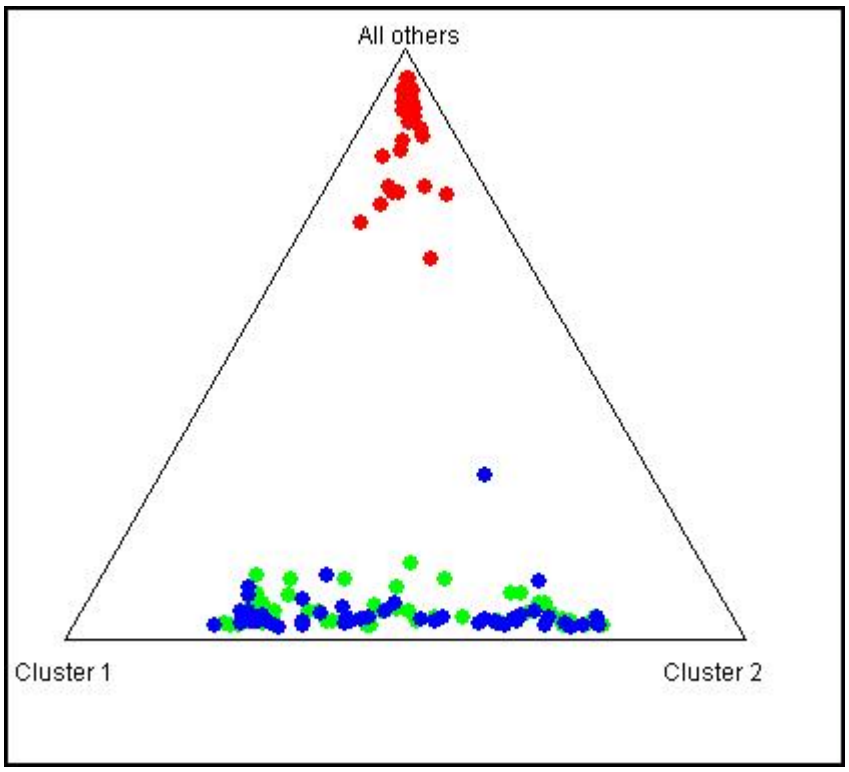


Figure 5. STRUCTURE results for regions (K=3) Florida (red), North Carolina (green) New York (blue).

Table 8. Estimated ln likelihoods of data for regions.

# groups (K)	Ln[P(X/K)]
1	-4283.7
2	-4062.7
3	-4041.5
4	-4276.9
5	-4352.5

Restoration Assessment

Allelic diversity was higher in the 2001 Anclote Estuary than the other samples for loci AICL112, AICL115 and AICL131; however, after correction for sample size, the allelic richness was comparable with that of the other samples. Deviation from Hardy-Weinberg equilibrium in the form of heterozygote deficit was observed in two samples (2001AN and 2005PI) for locus AICL112 and in two samples (2005PI and 1998ST) for locus AICL115 (Table 9). Linkage disequilibrium was observed for loci AICL112 and G340 in the 2001 Anclote sample. No other tests were significant.

Overall, microsatellite data indicated little to no population structure among the Florida populations sampled. All global F_{ST} values were less than or equal to 0.0025 and none were significant (Table 10). Exact tests for genic differentiation revealed no significant differences between populations, and charts of allelic frequencies for each locus showed a high degree of similarity among the populations (Appendix C). Using all nine loci, neither assignment tests nor Bayesian cluster analysis were able to differentiate among scallops originating at any three of these locations or between the two years sampled for Pine Island Sound. Due to similarity among populations, assignment tests correctly assigned individuals to their true population of origin less than 50% of the time (Figure 6, Table 11). Calculation of the average log likelihood (D_{LR}) in pairwise assignment comparisons showed a greater likelihood of individuals being assigned to populations other than their true source (-0.038 for 2001AN/2001PI, -0.026 for 2001AN/1998ST, and -0.046 for 2001PI/1998ST)

STRUCTURE results indicated that a single population was the most likely scenario for this group of samples. Visually this can be seen in the clustering of individual genotypes when

the number of groups was set at three (Figure 7), and there was no partitioning of the samples. In addition, the most likely (least negative log likelihood) number of populations was also one (Table 12). In WHICHLOCI analysis, use of all nine loci was incapable of achieving 95% correct assignment even at a stringency of Λ=0. Loci were ranked for their ability to assign individuals correctly with AICL131 and M26 most commonly first or second, followed by AICL115 as third. Loci S336 and AICL327 were generally the two lowest ranked loci.

As expected, given the inability of assignment to discriminate between Florida populations, assignment of the 50 individuals in the 2005PI sample in GENECLASS yielded no conclusive results. The sample was assigned relatively evenly among the three potential source populations, and few of these assignments had stringency above Λ=1. Analyses using exclusion probabilities were similarly inconclusive, as probabilities of exclusion were comparable among the three populations where either similarly high or low probabilities were calculated for all three populations. In all but a single case (using both assignment methods (Paetkau *et al*. (1995) and Rannala & Mountain (1997)), when 95% probability of exclusion was calculated for one population, at least 90% probability of exclusion was also calculated for the other two populations.

Table 9. Statistics of microsatellite loci in Florida samples Anclote (AN), Pine Island Sound (PI), and Steinhatchee (ST). Number of individuals scored (N), number of alleles observed (N_a), allelic richness (R) observed heterozygosity (H_o), expected heterozygosity (H_e), and *P*-values for heterozygote deficiency (*indicates significance at adjusted *P*-value).

Locus	Sample	N	N_a	R	H_o	H_e	*P*
AICL112	2001AN	50	21	14.366	0.760	0.863	0.0120*
	2001PI	50	12	9.981	0.800	0.831	0.2435
	2005PI	50	12	10.302	0.740	0.740	0.0148*
	1998ST	25	13	13.000	0.800	0.844	0.2435
AICL115	2001AN	50	12	8.972	0.600	0.707	0.0204
	2001PI	49	10	8.655	0.612	0.652	0.2000
	2005PI	49	8	7.542	0.469	0.685	0.0009*
	1998ST	25	7	7.000	0.400	0.614	0.0028*
AICL131	2001AN	49	16	12.598	0.614	0.818	0.0468
	2001PI	50	13	11.445	0.840	0.862	0.3037
	2005PI	49	13	10.662	0.857	0.849	0.5407
	1998ST	25	11	11.000	0.920	0.825	0.9417
AICL271	2001AN	50	9	8.445	0.840	0.788	0.8574
	2001PI	50	9	7.859	0.800	0.760	0.7796
	2005PI	50	8	6.990	0.860	0.747	0.9917
	1998ST	25	7	7.000	0.800	0.714	0.8787
AICL327	2001AN	50	6	4.876	0.300	0.303	0.5333
	2001PI	50	6	4.878	0.280	0.271	0.7157
	2005PI	50	6	4.630	0.300	0.285	0.7444
	1998ST	25	5	5.000	0.240	0.222	1.000
M26	2001AN	50	5	4.472	0.800	0.666	0.9898
	2001PI	50	7	5.884	0.780	0.710	0.8861
	2005PI	50	5	4.971	0.740	0.698	0.7787
	1998ST	25	6	6.000	0.720	0.666	0.7843
G340	2001AN	50	7	5.998	0.480	0.550	0.0630
	2001PI	50	5	4.988	0.660	0.593	0.9241
	2005PI	50	5	4.956	0.420	0.474	0.1083
	1998ST	25	7	7.000	0.720	0.629	0.9500
S336	2001AN	48	3	2.773	0.458	0.520	0.2056
	2001PI	49	3	2.510	0.388	0.483	0.0722
	2005PI	50	4	3.441	0.500	0.493	0.5685
	1998ST	25	3	3.000	0.560	0.495	0.7852
N391	2001AN	50	10	9.586	0.860	0.859	0.4778
	2001PI	47	10	9.661	0.809	0.873	0.0926
	2005PI	50	12	9.971	0.800	0.858	0.1046
	1998ST	25	8	8.000	0.720	0.861	0.0185

Table 10. F_{ST} (θ) and R_{ST} (ρ) coefficients between Florida samples Anclote Estuary (AN), Pine Island Sound (PI) and Steinhatchee (ST). F_{ST} values in upper right, R_{ST} values in lower left.

	2001AN	2001PI	2005PI	1998ST
2001AN		0.0022	0.0025	0.0014
2001PI	-0.0075		0.0019	-0.0007
2005PI	-0.0049	0.0009		0.0000
1998ST	-0.0074	-0.0001	0.0021	

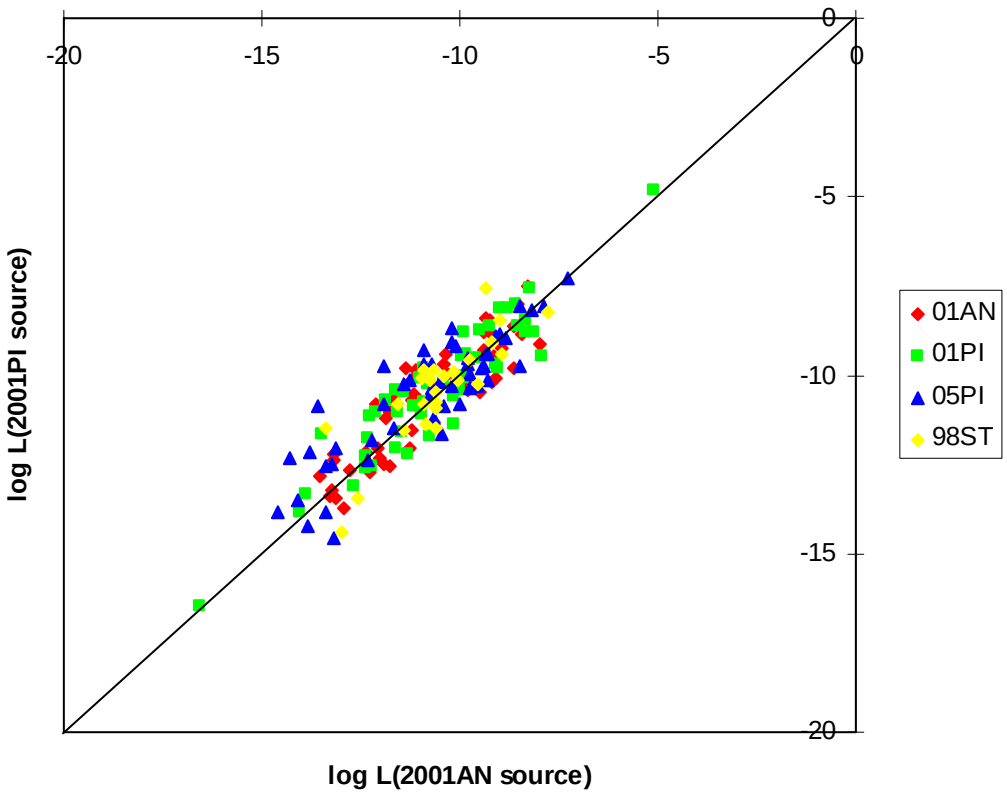


Figure 6. Assignment of Florida samples Anclote Estuary (AN), Pine Island Sound (PI) and Steinhatchee (ST) to 2001 Anclote Estuary and 2001 Pine Island Sound. Points to the right of the line indicate assignment to 2001AN, those to the left are assigned to 2001PI.

Table 11. Assignment of Florida samples Anclote Estuary (AN), Pine Island Sound (PI) and Steinhatchee (ST) to known source populations.

Source	Assigned to			% Correct
	01 AN	01 PI	98 ST	
01AN	13	22	15	26.0
01PI	14	23	13	46.0
98ST	9	8	8	32.0

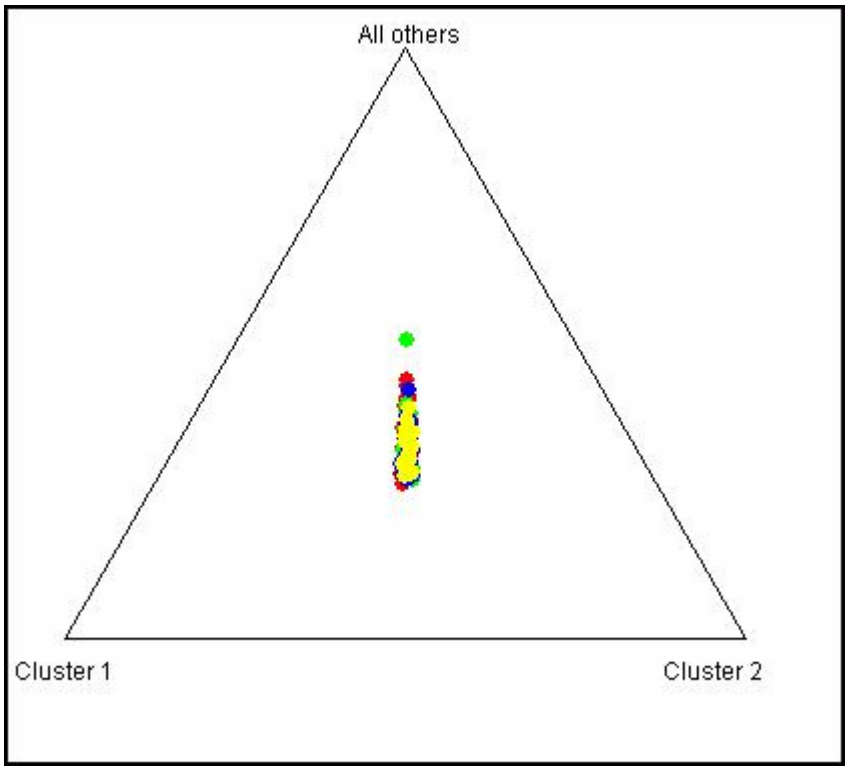


Figure 7. STRUCTURE results for Florida populations (K=3) 2001 Anclote Estuary (red), 2001 Pine Island Sound (green), 2005 Pine Island Sound (blue), and 1998 Steinhatchee (yellow).

Table 12. Estimated ln likelihoods of data for Florida populations.

# groups (K)	Ln[P(X/K)]
1	-4939.9
2	-4944.9
3	-5237.8
4	-5635.0
5	-5893.9

DISCUSSION

The regional analyses all supported strong differences between Florida and each of the Atlantic populations ($F_{ST} \geq 0.1137$, $R_{ST} \geq 0.0545$), though there is more agreement between exact tests of genic differentiation, assignment tests and Bayesian clustering with F_{ST} than with R_{ST}. According to F_{ST} (0.0354) and genic differentiation results, North Carolina and New York populations were much less differentiated than the Gulf and Atlantic. Results of R_{ST}, however, indicated similar levels of divergence between Florida and New York (0.0545) as between North Carolina and New York (0.0567), while divergence between Florida and North Carolina was much higher (0.1411). In comparing the F_{ST} and R_{ST} values, where R_{ST} is calculated using the Stepwise Mutation Model, one might propose that the lower R_{ST} between Florida and New York populations indicates that this divergence occurred more recently than between Florida and North Carolina. In this case, the divergences between New York and the other two populations would be less affected by mutations in the populations than the divergence between North Carolina and Florida. However, this idea does not appear to be supported by comparisons of allelic frequencies between New York and North Carolina, as both populations appear to contain similar allelic distributions (Appendix B). Therefore, the disagreement between F_{ST} and R_{ST} values may be due more to the high variance commonly observed in the R_{ST} statistic than to historical factors. Furthermore, results of all other tests, including assignment tests and Bayesian clustering conform more closely to the results of F_{ST} than R_{ST}.

Assignment tests were almost always able to identify the true source of individuals as either Florida or Atlantic; however, assignments were not as accurate between North Carolina and New York. Paetkau *et al.* (2004) found that power of assignment tests can be assessed using

the genotype likelihood ratio distance, D_{LR}, where a value of $D_{LR}>5$ indicates near maximum power of assignment, while a value less than three indicates low power of distinguishing immigrants from residents. The value of D_{LR} was close to five for assignment tests of Florida versus Atlantic populations, and this strongly supports the potential for application of assignment methods over this geographic area. However, a value of D_{LR} below one in assignments of New York and North Carolina, indicates low potential to use assignment tests for discriminating between scallops of these origins. Even though New York and North Carolina scallops were assigned correctly almost 75% of the time, the stringency of these assignments was very low. This supports the mtDNA work of Blake and Graves (1995), as well as the results of Bert *et al.* (in prep.), which has indicated most of the significant genetic variation in this species is distributed between Florida Gulf and Atlantic populations. However, it contradicts the current subspecies classification of *A. irradians* based on morphological data, which groups Florida and North Carolina populations together as *A.i. concentricus*, and considers the populations from Maryland and New Jersey to Cape Cod *A.i. irradians* (Blake & Shumway 2006).

A significant genetic break between the Gulf of Mexico and Atlantic populations is consistent with a pattern of divergence in the mitochondrial lineage of numerous species (Avise 1992). Among these species are marine, and coastal terrestrial animals with very different life-history characteristics, all sharing similar genetic breaks along the east coast of Florida reflecting separation of populations north and south of this boundary. Among this group of organisms is included the Eastern oyster (*Crassostrea virginica*), which shares similar characteristics of estuarine habitat, broadcast spawning and larval dispersal with the bay scallop. The explanation for this pattern is the geographic history of the southeastern U.S., which has undergone dramatic alterations in landscape with episodes of sea level rise and fall during the Pleistocene.

Specifically, changes in the size and shape of the Florida peninsula have alternately caused expansion and contraction of coastal habitats such as estuaries and salt marshes (Avise 1992). Additional temperature changes associated with global warming and cooling would have shifted the ranges of tropical and temperate adapted organisms, at times turning the Florida peninsula into a geographic barrier preventing dispersal of organisms adapted to cooler climates.

Modern oceanic circulation patterns could be maintaining relative isolation of the Gulf from the Atlantic while causing some migration between North Carolina and New York with larval dispersal by northward moving currents. The similarity between the Atlantic populations may indicate current high levels of naturally occurring gene flow between North Carolina and New York, or, as suggested by Rhodes (1991), a significant amount of mixing of these populations mediated by aquacultural activities. However, given the structure observed among these regions begs the question of how appropriate the current classification of this organism is. These data indicate there is a Gulf of Mexico assemblage separated genetically from an Atlantic assemblage, and thus the current *A.i. concentricus* and *A.i.irradians* subspecies lack support with these genetic markers.

While the regional study demonstrated that this set of microsatellite markers is capable of identifying differences between populations and identifying individuals originating from different sources, there was not sufficient differentiation between Florida populations to successfully assign individuals within that region. The purpose of this case study has been to determine whether there has been a recent influx of immigrant genotypes (restoration scallops) into a population using assignment tests. To apply this method, immigrant genotypes must be distinct from native genotypes, and therefore there must be some differentiation between source and sink populations. If there is no genetic differentiation between two populations, immigrants

can not be identified because migration is occurring naturally at a high enough rate to homogenize the gene pools. The results of the analyses comparing the Florida populations revealed little or no difference between populations ($F_{ST} \leq 0.0025$, $R_{ST} \leq 0.0021$), possibly indicating high levels of naturally occurring gene flow between Anclote Estuary, Pine Island Sound and Steinhatchee. Assignment tests and Bayesian cluster analysis all failed to confidently identify scallops as more likely to originate in any one of these three populations. Values of D_{LR} close to zero indicates almost no power of assignment tests to distinguish between scallops originating in these Florida populations. Furthermore, the 2005 Pine Island Sound post-restoration sample yielded no indication of genetic differentiation from any of the other populations, nor could these individuals be assigned or excluded from any source population with confidence. However, Paetkau *et al.* (2004) notes that the power of assignment tests increases with the number of loci, therefore, additional loci could increase the power of the assignments.

In addition, there is a large amount of variation in the ability for individual loci to reveal population structure or to identify an individual scallop as a member of one population or another. As noted for the regional analyses, a single locus (AICL327) is responsible to a large degree for the ability to differentiate between Florida and Atlantic scallops. It is possible that other loci will demonstrate more isolation of these Gulf populations than those used here. Lack of population structure in the region is inconsistent with what is known about recruitment from a study by Arnold *et al.* (1998), which inferred from scallop abundance data in consecutive years that recruitment to Gulf populations (including Anclote Estuary and Steinhatchee) was largely localized. Therefore, there may be sufficient gene flow to keep these populations genetically similar, but not to noticeably affect abundance.

Barring assignment tests and Bayesian cluster analysis to determine the impact of the 2003 Pine Island Sound restoration, we can look at other genetic signatures that may result from a sudden influx of larvae produced from a small number of individuals. The restoration broodstock used to produce the larvae for this restoration was comprised of 12 scallops from Anclote Estuary, contributing a maximum of 24 alleles. As spawning bay scallops may not contribute equally to the offspring resulting in a dramatic effect of genetic drift in the offspring (Hedgecock *et al.* 1992), as observed in chapter one, the maximum number of alleles of the released larvae (F_1 offspring) ought to be much lower than that of a naturally occurring bay scallop cohort (Evans *et al.* 2004). If the wild population is indeed a small number of individuals (as indicated by observations of low abundance), given a large influx of the released larvae, one might expect to observe a detectable shift in allele frequencies, heterozygosity, linkage disequilibrium and/or effective population size in the 2005PI sample relative to the 2001PI baseline population. However, we did not observe a consistent decrease in number of alleles or heterozygosity in the 2005 sample compared to 2001. The relatively similar genetic profile of the 2005 and 2001 Pine Island Sound populations, as well as the similarity between Pine Island Sound, Anclote Estuary and Steinhatchee, indicates that these disjunct populations are functioning as an assemblage with at least some gene flow and are likely to be supplementing each other through migration of larvae.

The findings of this study indicate that the method of restoration used in this case, where genetic samples of the restoration stock and broodstock are not available, is not compatible with an evaluation using assignment tests based on wild source populations that are not genetically distinct. However, given genotypes from samples of the broodstock and restoration stock used for the restoration may provide a more accurate baseline sample with which to conduct these

analyses. Based on the findings of chapter one, as well as previous observations of unequal contribution of shellfish in aquaculture spawns, the strong effect of genetic drift can dramatically alter the gene pool of the restoration sample. Given sufficient genetic differentiation between released larvae and the wild population of scallops (e.g. F_{ST}~0.1), assignment tests may be effective.

LITERATURE CITED

Arnold, W.S. 2001. Bivalve enhancement and restoration strategies in Florida, U.S.A. *Hydrobiologia* 465:7-19.

Arnold, W.S., D.C. Marelli, C.P. Bray & M.M. Harrison. 1998. Recruitment of bay scallops *Argopecten irradians* in Florida Gulf of Mexico waters: scales of coherence. *Mol. Ecol. Prog. Ser.* 170:142-157.

Avise, J.C. 1992. Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* 63: 62-76.

Balloux, F. & N. Lugon-Moulin. 2002. The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11:155-165.

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

Barber, B.J. & N.J. Blake. 2006. Reproductive Physiology. In S.E. Shumway and G.J. Parsons (Eds). Scallops: Biology,Ecology and Aquaculture 2nd Ed., Developments in Aquaculture and Fisheries Science, Vol. 35. Elsevier B.V., Amsterdam. pp. 357-416.

Bass, A.L., S.P. Epperly & J. Braun-McNeill. 2004. Multi-year analysis of stock composition of a loggerhead turtle (*Caretta caretta*) foraging habitat using maximum likelihood and Bayesian methods. *Conservation Genetics* 5: 783-796.

Beacham, T.D., K.D. Le & J.R. Candy. 2004. Population structure and stock identification of steelhead trout (*Oncorhynchus mykiss*) in British Columbia and the Columbia River based on microsatellite variation. *Environ. Biol. Fish.* 69: 95-109.

Beacham, T.D. & C.C. Wood. 1999. Application of microsatellite DNA variation to estimation of stock composition and escapement of Nass River sockeye salmon (*Oncorhynchus nerka*). *Can. J. Fish. Aquat. Sci.* 56(2): 297-310.

Bernatchez, L. & P. Duchesne. 2000. Individual-based genotype analysis in studies of parentage and population assignment: how many loci, how many alleles? *Can. J. Fish. Aquat. Sci.* 57: 1-12.

Berry, O., M.D. Tocher & S.D. Sarre. 2004. Can assignment tests measure dispersal? *Mol. Ecol.* 13:551-561.

Blake, S.G. & J.E. Graves. 1995. Mitochondrial DNA variation in the bay scallop, *Argopecten irradians* (Lamarck, 1819), and the Atlantic calico scallop, *Argopecten gibbus* (Linnaeus, 1758). *J. Shellfish Res.* 14:79-85.

Blake, N.J. & S.E. Shumway. 2006. Bay Scallop and Calico Scallop Fisheries, Culture and Enhancement in Eastern North America, . In S.E. Shumway and G.J. Parsons (Eds). Scallops: Biology,Ecology and Aquaculture 2nd Ed., Developments in Aquaculture and Fisheries Science, Vol. 35. Elsevier B.V., Amsterdam. pp. 945-964.

Boudry, P., B. Collet, F. Cornette, V. Hervouet & F. Bonhomme. 2002. High variance in reproductive success of the Pacific oyster (*Crassostrea gigas*, Thornberg) revealed by microsatellite-based parentage analysis of multifactorial crosses. *Aquaculture* 204: 283-296.

Bologna, P.A.X. 1998. Growth, production, and reproduction in bay scallops *Argopecten irradians concentricus* (Say) from the northern Gulf of Mexico. *J. Shellfish Res*. 17(4): 911-917.

Bologna, P.A.X., A.E. Wilbur & K.W. Able. 2001. Reproduction, population structure, and recruitment limitation in a bay scallop (*Argopecten irradians* Lamarck) population from New Jersey, USA. *J. Shellfish Res*. 20(1):89-96.

Bravington, M. V. & R. D. Ward. 2004. Microsatellite DNA markers: evaluating their potential for estimating the proportion of hatchery-reared offspring in a stock enhancement programme. *Mol. Ecol.* 13: 1287-1297.

Bricelj, V.M. & M.K. Krause. 1992. Resource allocation and population genetics of the bay scallop, *Argopecten irradians irradians*: effects of age and allozyme heterozygosity on reproductive output. *Mar. Biol*. 113(2):253-261.

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

Brownstein, M.J., J.D. Carpten & J.R. Smith. 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* 20(6): 1004-6, 1008-10.

Burton, R.S. & M.J. Tegner. 2000. Enhancement of red abalone *Haliotis rufescens* stocks at San Miguel Island: reassessing a success story. *Mar. Ecol. Prog. Ser.* 202: 303-308.

Chikarmane, H.M., A.M Kuzirian, I. Carroll, & R. Dengler. 2001. Development of genetically tagged bay scallops for evaluation of seeding programs. *Biol. Bull*. 201:285-286.

Cornuet, J.-M., S. Piry, G. Luikart, A. Estoup & M. Solignac. 1999. New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* 153: 1989-2000.

Eckman, J.E. 1987. The role of hydrodynamics in recruitment, growth, and survival of *Argopecten irradians* (L.) and *Anomia simplex* (D'Orbigny) within eelgrass meadows. *J. Exp. Mar. Biol. Ecol*. 106: 165-191.

Eldridge, M.D.B., J.E. Kinnear & M.L. Onus. 2001. Source population of dispersing rock-wallabies (*Petrogale lateralis*) identified by assignment tests on multilocus genotypic data. *Mol. Ecol*. 10:2867-2876.

El Mousadik, A. & R.J. Petit. 1996. Chloroplast DNA phylogeography of the argan tree of Morocco. *Mol. Ecol.* 5:547-555.

Epifanio J.M., P.E. Smouse, C.J. Kobak & B.L. Brown. 1995. Mitochondrial DNA divergence among populations of American shad (*Alosa sapidissima*): how much variation is enough for mixed-stock analysis? *Can. J. Fish. Aquat. Sci* 52:1688-1702.

Evans, B., J. Bartlett, N. Sweijd, P. Cook, & N.G. Elliot. 2004. Loss of genetic variation at microsatellite loci in hatchery produced abalone in Australia (*Haliotis rubra*) and South Africa (*Haliotis midae*). *Aquaculture* 233: 109-127.

Farias, A. & I. Uriarte. 2006. Nutrition in Pectinids . In S.E. Shumway and G.J. Parsons (Eds). *Scallops: Biology,Ecology and Aquaculture* 2nd Ed., Developments in Aquaculture and Fisheries Science, Vol. 35. Elsevier B.V., Amsterdam. pp. 521-542.

Frankham, R., J.D. Ballou & D.A. Briscoe. 2002. Introduction to Conservation Genetics. Cambridge, United Kingdom: Cambridge University Press.

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

Gaffney, P.M., V.P. Rubin, D. Hedgecock, D.A. Poers, G. Morris & L. Hereford. 1996. Genetic effects of artificial propagation – signals from wild and hatchery populations of red abalone in California. *Aquaculture* 143: 257-266.

Gjetvaj, B., R.M. Ball, S. Burbridge, C.J. Bird, E. Kenchington & E. Zouros. 1997. Amounts of polymorphism at microsatellite loci in the sea scallop *Placopecten magellanicus*. *J. Shellfish Res*. 16(2): 547-553.

Goldberg, R., J. Pereira & P. Clark. 2000. Strategies for enhancement of natural bay scallop, *Argopecten irradians irradians*, populations; a case study in the Niantic River estuary, Connecticut, USA. *Aquaculture Int*. 8: 139-158.

Goudet, J. 1995. FSTAT Version 1.2: a computer program to calculate F-statistics. *J. Hered.* 86:485-486

Hansen, M.M., E. Kenchington & E.E. Nielsen. 2001. Assigning individual fish to populations using microsatellite DNA markers. *Fish and Fisheries* 2:93-112.

Hare, M.P., S.A. Karl, & J.C. Avise. 1996. Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. *Mol. Biol. Evol.* 13(2):334-345.

Hare, M.P., S.K. Allen, Jr., P. Bloomer, M.D. Camara, R.B. Carnegie, J. Murfree, *et al.* 2005. A genetic test for recruitment enhancement in Chesapeake Bay oysters, *Crassostrea virginica*, after population supplementation with a disease tolerant strain. *Con. Genet.* 7(5):717-734.

Hayes, B., A.K. Sonesson & B. Gjerde. 2005. Evaluation of three strategies using DNA markers for traceability in aquaculture species. *Aquaculture* 250:70-81.

Hedgecock, D. & F. Sly. 1990. Genetic drift and effective population sizes of hatchery propagated stocks of the Pacific oyster, *Crassostrea gigas. Aquaculture* 88:21-38.

Hedgecock, D., V. Chow & R.S. Waples. 1992. Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies. *Aquaculture* 108: 215-232.

Hedgecock, D., G. Li, S. Hubert, K. Bucklin & V. Ribes. 2004. Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the pacific oyster, *Crassostrea gigas. J. Shellfish Res.* 23(2):379-385.

Hoarau, G., A.M.-T. Piquet, H.W. van der Veer, A.D. Rijnsdorp, W.T. Stam & J.L. Olsen. 2004. Population structure of plaice (*Pleuronectes platessa* L.) in northern Europe: a comparison of resolving power between microsatellites and mitochondrial DNA data. *J. Sea Res*. 51: 183-190.

Hoeh, W.R., D.T. Stewart & S.I. Guttman. 2002. High fidelity of mitochondrial genome transmission under the doubly uniparental mode of inheritance in freshwater mussels (Bivalvia: Unionoidea). *Evolution* 56(11): 2252-2261.

Jarne, P. & P.J.L. Lagoda. 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol*. 11:424-429.

Kaukinen, K.H., K.J. Supernault & K.M. Miller. 2004. Enrichment of tetranucleotide microsatellite loci from invertebrate species. *J. Shellfish Res*. 23(2):621-626.

Kim, J. E., R. E. Withler, C. Ritland & K. M. Cheng. 2004. Genetic variation within and between domesticated chinook salmon, *Oncorhynchus tshawytscha*, strains and their progenitor populations. *Env. Biol. Fish.* 69: 371-378.

Kimura, M. & J.F. Crow. 1964. The number of alleles that can be maintained in a finite populations. *Genetics* 49:725-738.

Kimura, M. & T. Ohta. 1978. Stepwise muatation model and distribution of allelic frequencies in a finite populations. *Proc. Nat. Acad. Sci.USA*. 75:2868-2872.

Knowlton, N. & J.B.C. Jackson. 1993. Inbreeding and outbreeding in marine invertebrates. In: *The Natural History of Inbreeding and Outbreeding* (ed. N.W. Thornhill), Chicago, University of Chicago Press, pp 200-249.

Kocher, T.D., W.K. Thomas, A. Meyer, S.V. Edwards, S. Paabo, F.X. Villablanca, & A.C. Wilson. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proc.Nat. Acad. Sci.* 86(16):6196-6200.

Krause, M.K. 1992. Use of genetic markers to evaluate the success of transplanted bay scallops. *J. Shellfish Res*. 11:199.

Kreitman, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. *Nature* 304(4): 412-417.

Launey, S., M. Barre, A. Gerard & Y. Naciri-Graven. 2001. Population bottleneck and effective size in *Bonaima ostreae*-resistant populations of *Ostrea edulis* as inferred by microsatellite markers. *Genet. Res.* 78: 259-270.

Leverone, J.R. 2004. Restoration of bay scallop populations in Pine Island Sound: competent larval release strategy. Draft Final Report. Mote Marine Laboratory: 13 pp.

Liao, I.C., M.S. Su & E.M. Leano. 2003. Status of research in stock enhancement and sea ranching. *Rev. Fish Biol. Fish.* 13: 151-163.

Manel, S., O.E. Gaggiotti & R.S. Waples. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol*. 20(3):136-142.

Martinez, J.L., J. Dumas, E. Beall & E. Garcia-Vazquez. 2001. Assessing introgression of foreign strains in wild Atlantic salmon populations: variation in microsatellites assessed in historic scale collections. *Freshwater Biol.* 46: 835-844.

Michalakis, Y. & L. Excoffier. 1996. A generic estimation of population subdivision using distances between alleles with special reference to microsatellite loci. *Genetics* 142:1061-1064.

Milbury, C.A., D.W. Meritt, R.I.E. Newell & P.M. Gaffney. 2004. Mitochondrial DNA markers allow monitoring of oyster stock enhancement in the Chesapeake Bay. *Mar. Biol.* 145: 351-359.

Paekall, R. & P.E. Smouse. 2005. GenAlEx 6: Genetic Analysis in Excel. Population genetic software for teaching and research. The Australian National University, Canberra, Australia. http://www.anu.edu.au/BoZo/GenAlEx/

Paetkau, D., W. Culvert, I. Stirling & C. Strobeck. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* 4: 347-354.

Paetkau, D., L.P. Waits, P.L. Clarkson, L. Craighead & C. Strobeck. 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics* 147:1943-1957.

Paetkau, D., R. Slade, M. Burden & A. Estoup. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* 13: 55-65.

Peterson, C.H. & H.C. Summerson. 1992. Basin-scale coherence of population dynamics of an exploited marine invertebrate, the bay scallop: implications of recruitment limitation. *Mar. Ecol. Prog. Ser*. 90:257-272.

Peterson CH, HC Summerson, RA Luettich Jr. 1996. Response of bay scallops to spawner transplants: a test of recruitment limitation. *Mar. Ecol. Prog. Ser*. 132: 93-107.

Piry, S., A. Alapetite, J.-M. Cornuet, D. Paetkau, L. Baudouin & A. Estoup. 2004. GENECLASS2: A software for genetic assignment and first-generation migrant detection. *J. Hered.* 95(6): 536-539.

Potvin, C. & L. Bernatchez. 2001. Lacustrine spatial distribution of landlocked Atlantic salmon populations assessed across generations by multilocus individual assignment and mixed-stock analyses. *Mol. Ecol.* 10:2375-2388.

Primmer, C.R., M.T. Koskinen & J. Pironen. 2000. The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud. *Proc. R. Soc. Lond. B* 267: 1699-1704.

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

Pritchard, J.K. & W. Wen. 2004. Documentation for *structure* software: Version 2. University of Chicago. http://pritch.bsd.uchicago.edu.

Queller, D.C., J.E. Strassmann & C.R. Hughes. 1993. Microsatellites and Kinship. *Trends Ecol. Evol.* 8(88): 285-288.

Rannala, B. & J.L.Mountain. 1997. Detecting immigration by using multilocus genotypes. *Proc. Nat. Acad. Sci. USA.* 94: 9197-9201.

Raymond, M. & F. Rousset. 1995. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *J. Heredity* 86: 248-249.

Reynolds, J.H. & W.D. Templin. 2004. Detecting specific population in mixtures. *Env. Biol. Fish*. 69: 233-243.

Rhodes, E.W. 1991. Fisheries and aquaculture of the bay scallop, *Argopecten irradians*, in the eastern United States. In: *Scallops: Biology, Ecology and Aquaculture* (ed. S.E. Shumway), Elsevier, Amsterdam, pp 913-924.

Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution* 43(1): 223-225.

Roberts, S., C. Romano & G. Gerlach. 2005. Characterization of EST derived SSRs from the bay scallop, *Argopecten irradians*. *Mol. Ecol. Notes* 5(3): 567-568.

Rousset, F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* 142: 1357-1362.

Ruzzante, D.E., M.M. Hansen, D. Meldrup & K.M. Ebert. 2004. Stocking impact and migration pattern in an anadromous brown trout (*Salmo trutta*) complex: where have all the stocked spawning sea trout gone? *Mol. Ecol.* 13: 1433-1445.

Saavedra, C. 1997. Low effective sizes in hatchery population of the European oyster (*Ostrea edulis*): implications for the management of genetic resources. *J.Shellfish Res.* 16(2):441-446.

Shaw, P.W., C. Turan, J.M. Wright, M. O'Connell, & G.R. Carvalho. 1999. Microsatellite DNA analysis of population structure in Atlantic herring (*Clupea harengus*), with direct comparison to allozyme and mtDNA RFLP analyses. *Heredity* 83: 490-499.

Smouse, P.E., C.J. Kobak & S. Xu. 1994. Some thought on information content in allozyme and DNA markers in genetic stock identification. In: Application of DNA Technology to the Management of Pacific Salmon, L.K. Park, P.C. Moran & R.S. Waples. U.S. Dept. Commerce, NOAA Tech Memo NMFS-NWFSC-17. pp. 121-126.

Tettelbach, S.T. & P. Wenzcel. 1993. Reseeding efforts and the status of bay scallop *Argopecten* irradians (Lamarck, 1819) populations in New York following the occurrence of "brown tide" algal blooms. *J. Shellfish Res*. 12(2):423-431.

Thayer, G.W. & H.H. Stuart. 1974. The bay scallop makes its bed of seagrass. *Mar.Fish. Rev.* 35: 27-30.

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

Vasemagi, A., R. Gross, T. Paaver, M. Kangur, J. Nilsson & L.-O. Eriksson. 2001. Identification of the origin of an Atlantic salmon (*Salmo salar* L.) population in a recently recolonized river in the Baltic sea. *Mol. Ecol.* 10: 2877-2882.

Waser, P.M. & C. Stroebeck. 1998. Genetic signatures of interpopulation dispersal. *Trends Ecol. Evol.* 13(2) 43-44.

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

Wilbur, A.E. 1995. Population Genetics of the Bay Scallop, *Argopecten irradians* (Lamarck): An Analysis of geographic variation and the consequences of self-fertilization. PhD dissertation, University of Delaware, Newark.

Wilbur, A.E., S. Seyoum, T.M. Bert & W.S. Arnold. 2005. A genetic assessment of bay scallop (*Argopecten irradians*) restoration efforts in Florida's Gulf of Mexico coastal waters (USA). *Con. Genet.* 6(1):111-122.

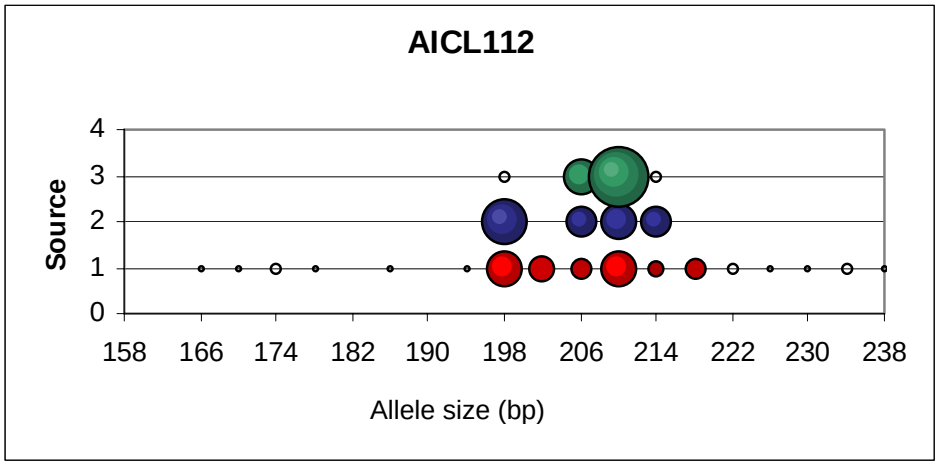
Zane, L., L. Bargelloni & T. Patarnello. 2002. Strategies for microsatellite isolation: a review. *Mol. Ecol.* 11:1-16.

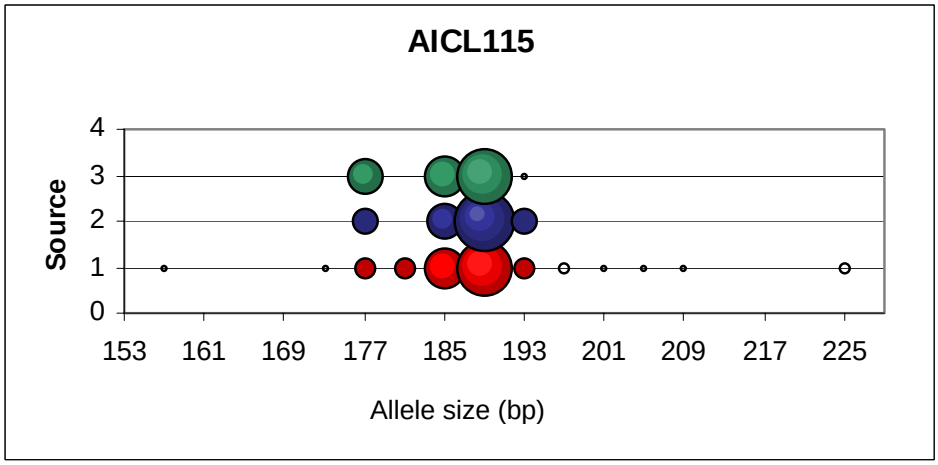
Zar, J.H. 1999. *Biostatistical Analysis* 4th Ed. Delhi, India: Pearson Education.

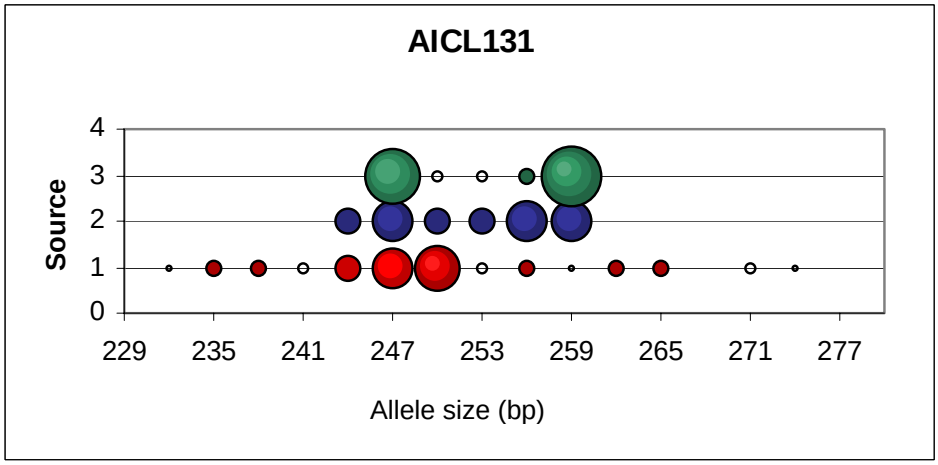
Zhan, A.-B., Z.M. Bao, X.L. Wang & J.J. Hu. 2005. Microsatellite markers derived from bay scallops *Argopecten irradians* expressed sequence tags. *Fish. Sci.* 71:41-1346.

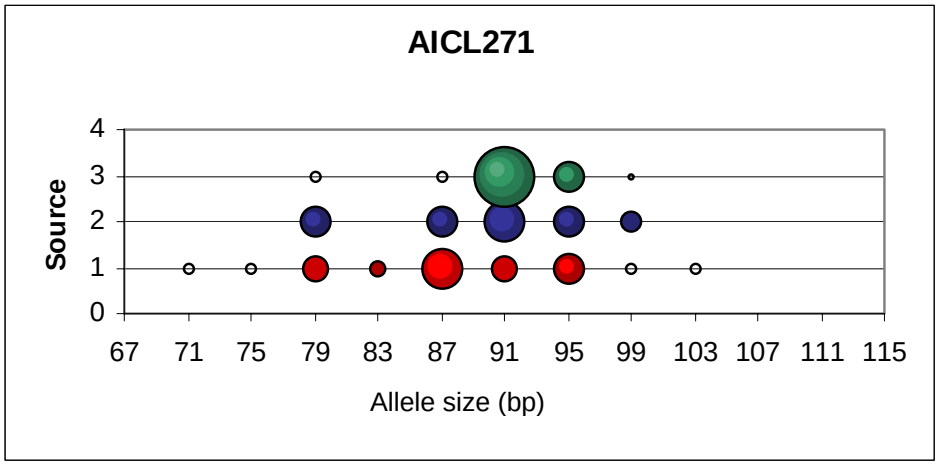
APPENDIX A

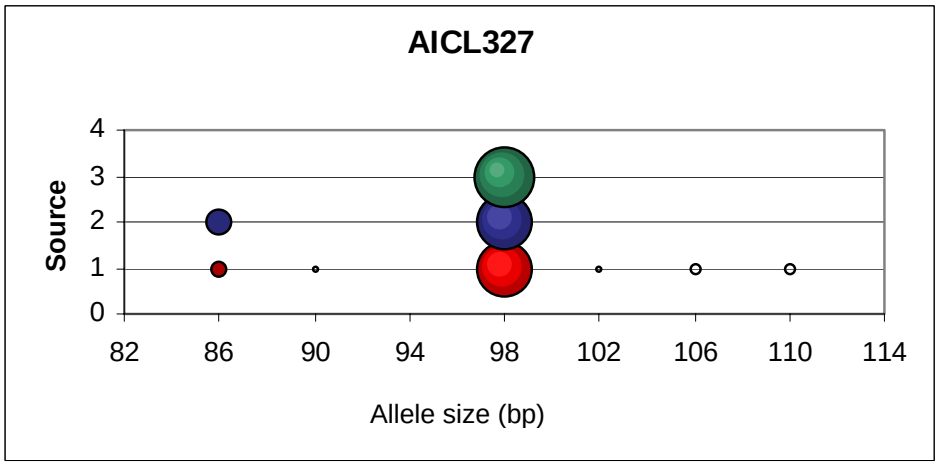
Allele frequencies of *A. irradians* microsatellites in (1) 2001 Anclote Estuary wild sample, (2) 04psk parents, and (3) 04byop offspring.





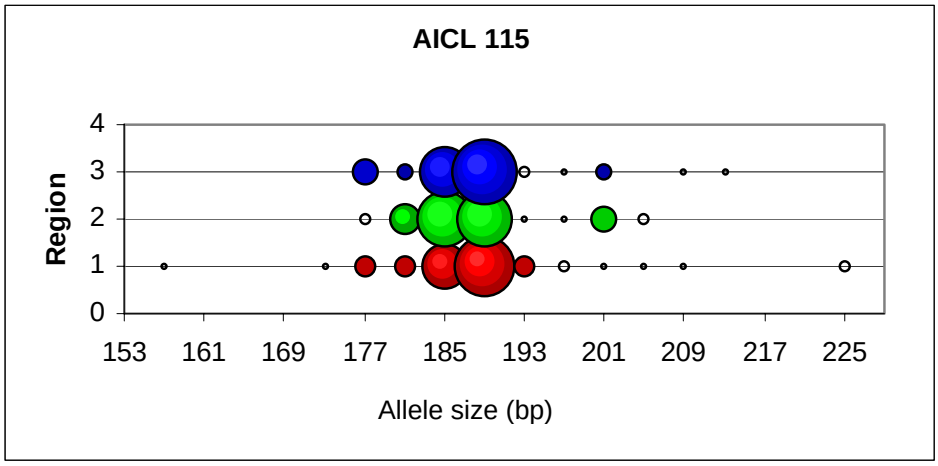


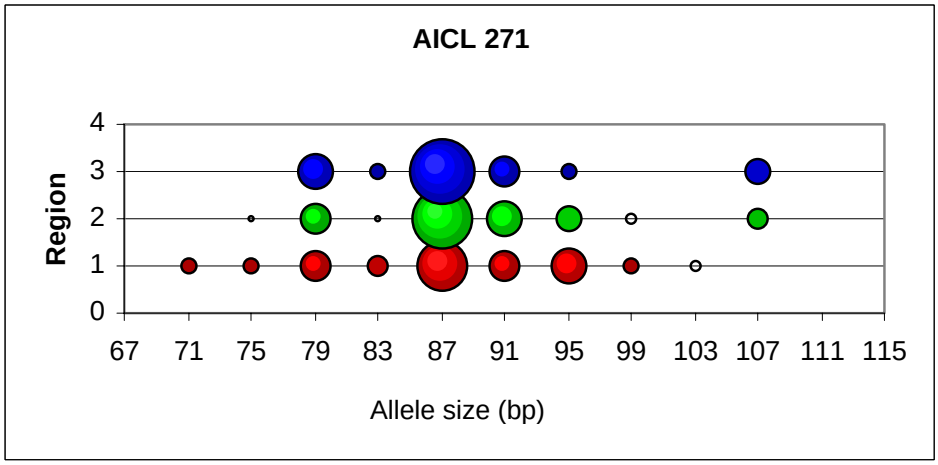

AICL271
Source
4
3
2
1
0
67
71
75
79
83
87
91
95
99
103
107
111
115
Allele size (bp)

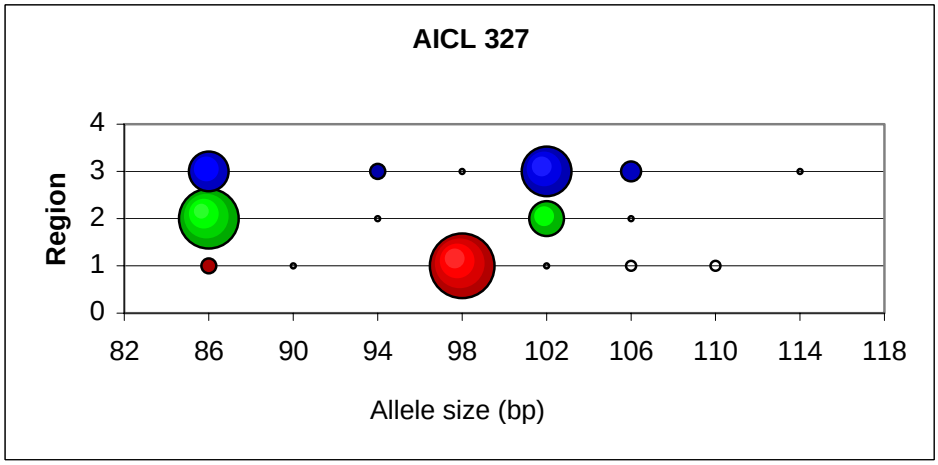

AICL327
Source
4
3
2
1
0
82
86
90
94
98
102
106
110
114
Allele size (bp)

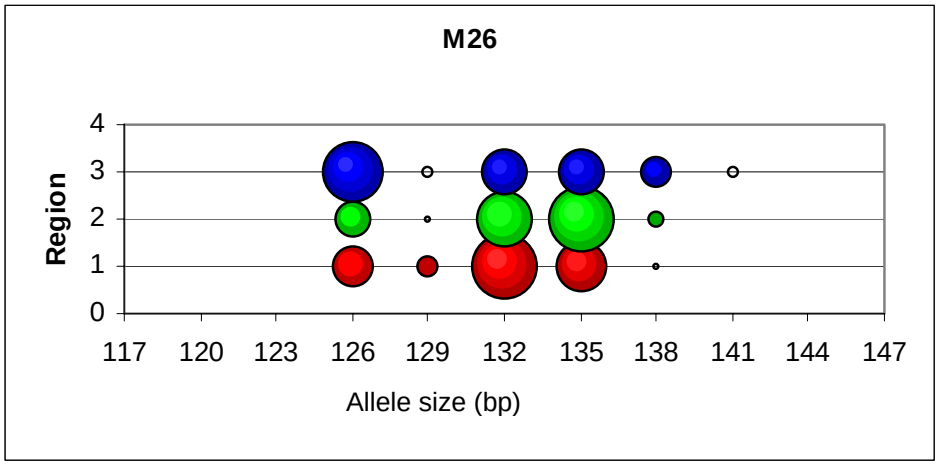
APPENDIX B

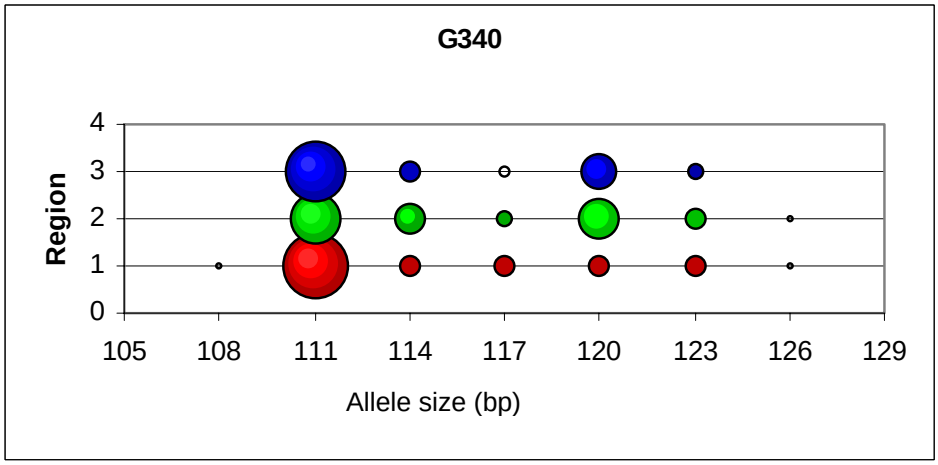
Allele frequencies of *A. irradians* microsatellites in regional populations (1) Florida, (2) North Carolina, and(3) New York.

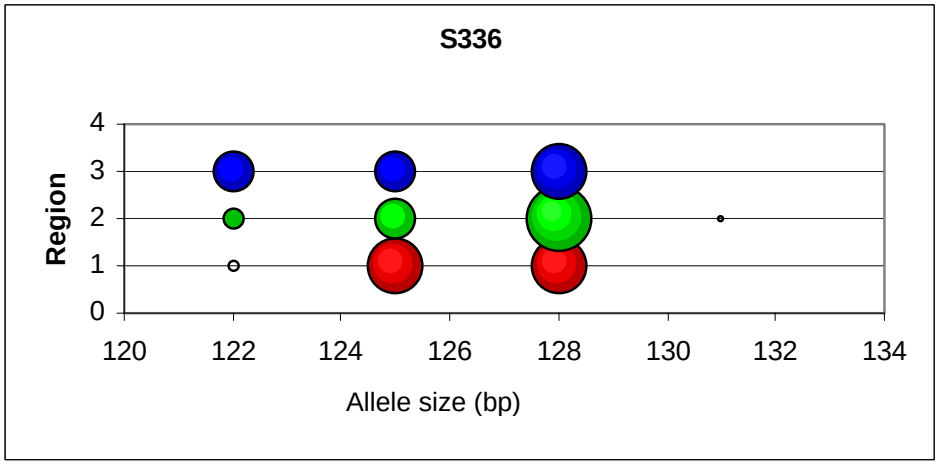


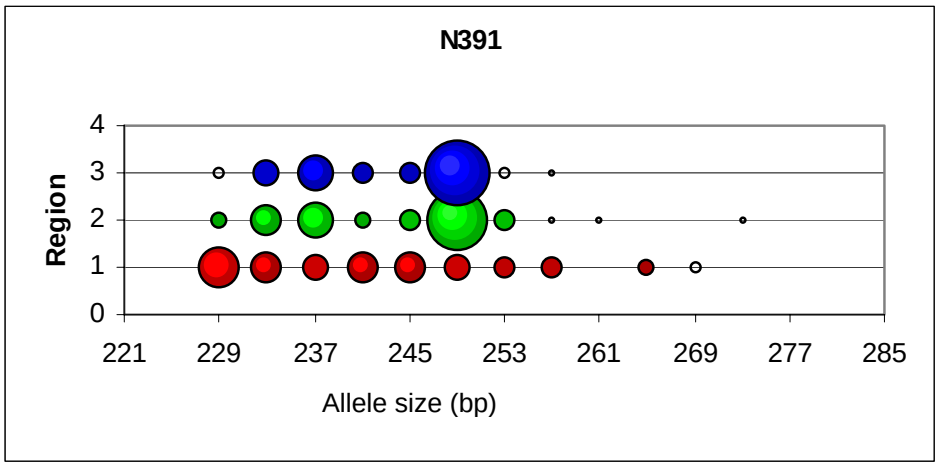


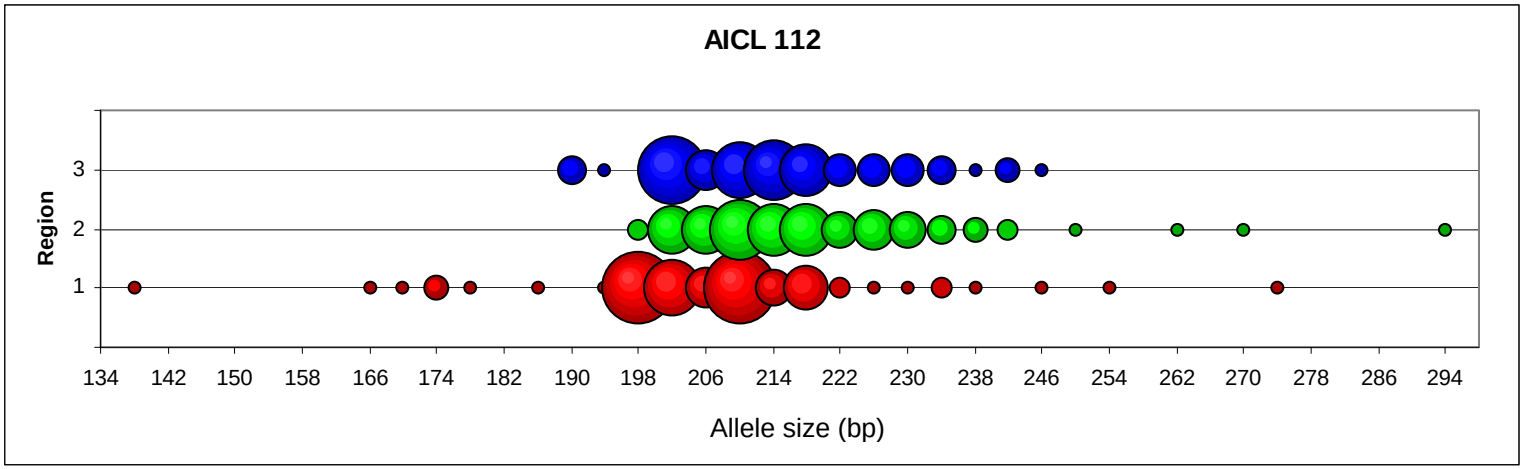


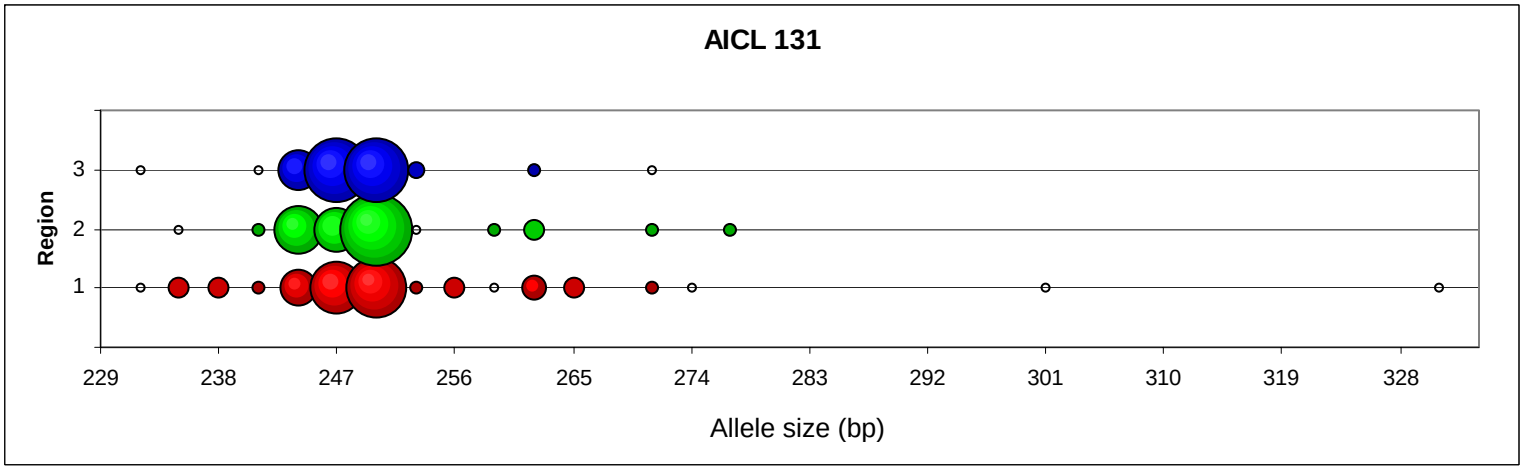

M26
Region
4
3
2
1
0
117
120
123
126
129
132
135
138
141
144
147
Allele size (bp)


G340
Region
4
3
2
1
0
105
108
111
114
117
120
123
126
129
Allele size (bp)


S336
Region
4
3
2
1
0
120
122
124
126
128
130
132
134
Allele size (bp)

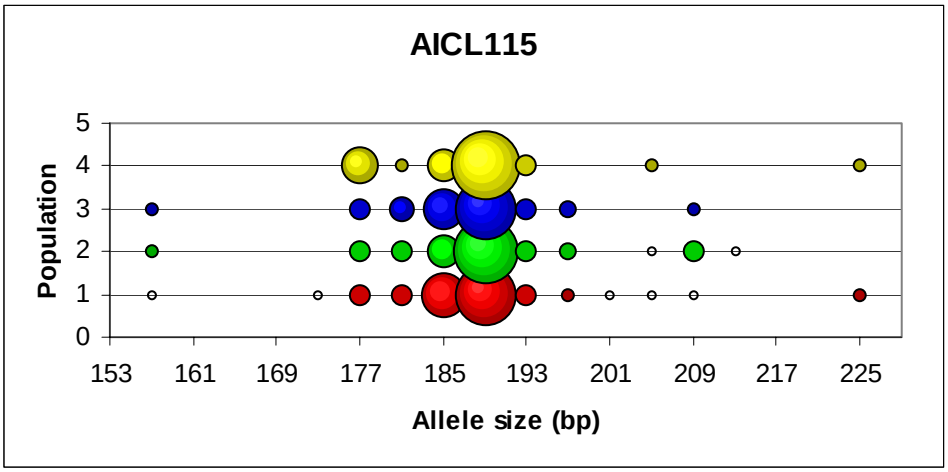

N391
Region
4
3
2
1
0
221
229
237
245
253
261
269
277
285
Allele size (bp)

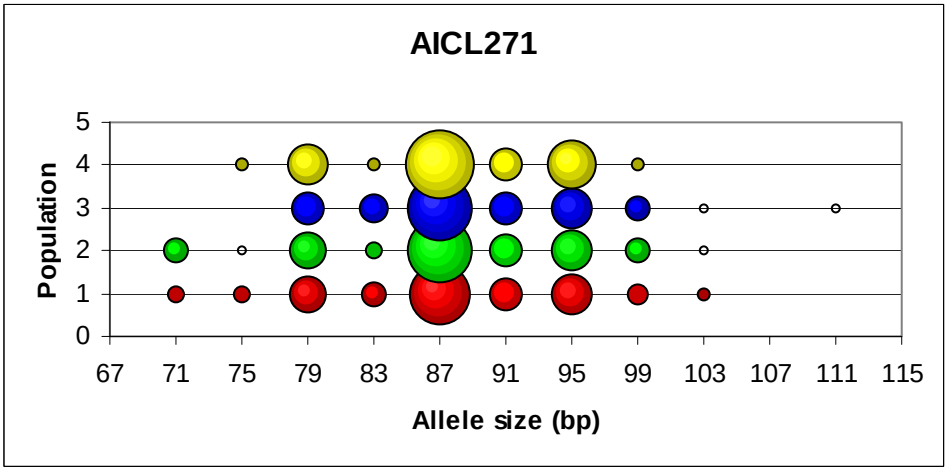

AICL 112
Region
3
2
1
134 142 150 158 166 174 182 190 198 206 214 222 230 238 246 254 262 270 278 286 294
Allele size (bp)

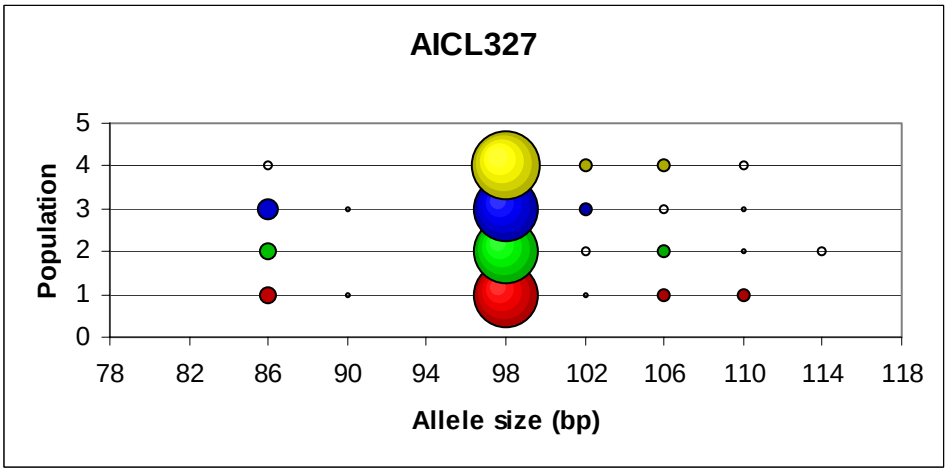

AICL 131
Region
3
2
1
229 238 247 256 265 274 283 292 301 310 319 328
Allele size (bp)

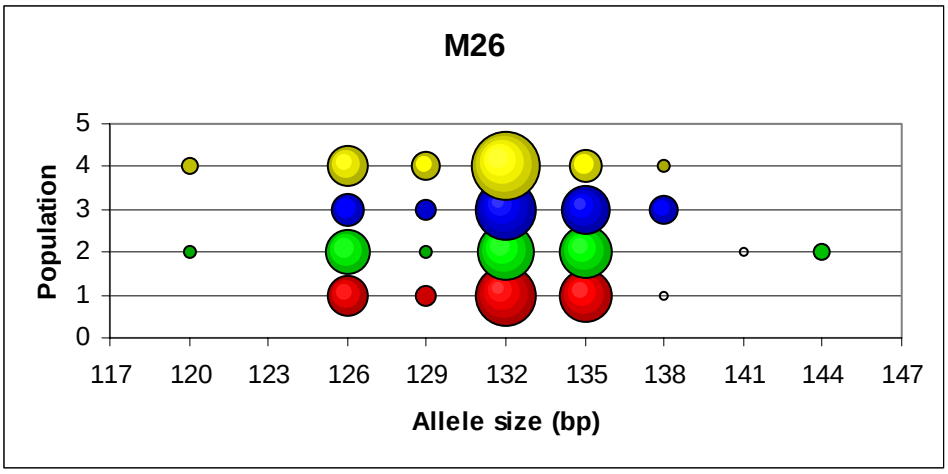
APPENDIX C

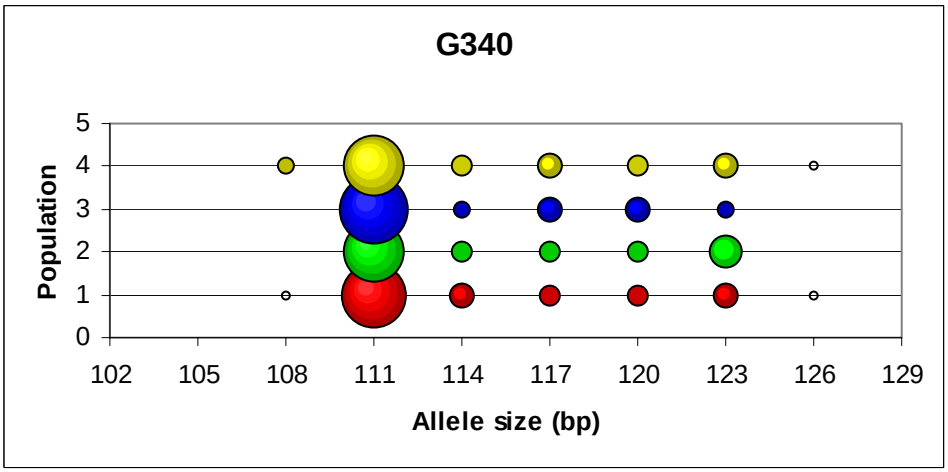
Allele frequencies of *A. irradians* microsatellites in Florida populations
(1) 2001 Anclote Estuary, (2) 2001 Pine Island Sound,
(3) 2005 Pine Island Sound, and (4) 1998 Steinhatchee.

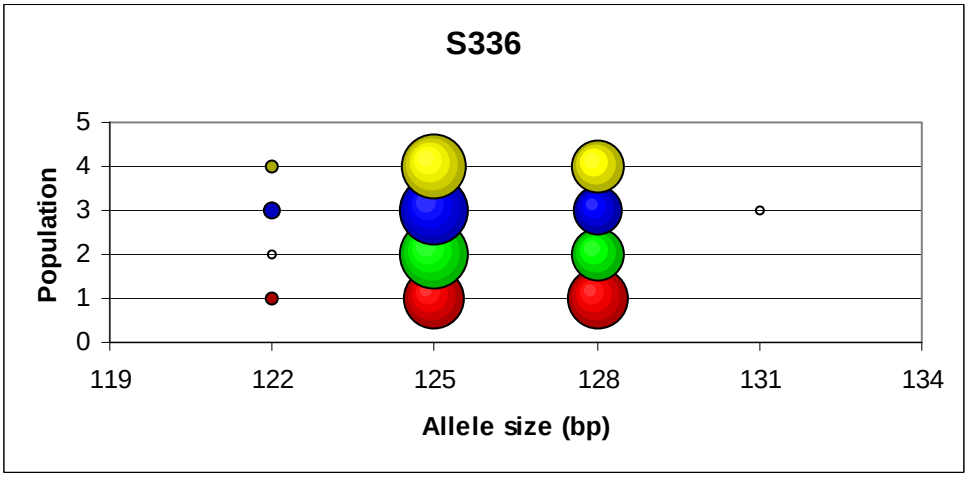


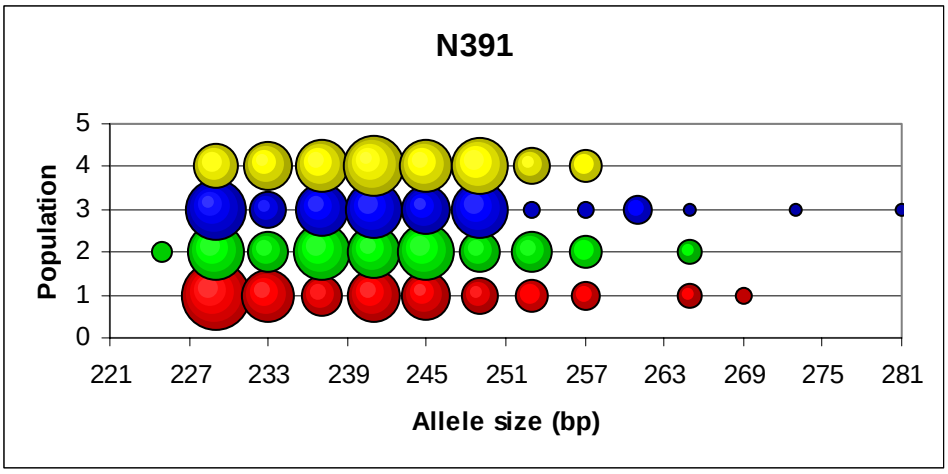


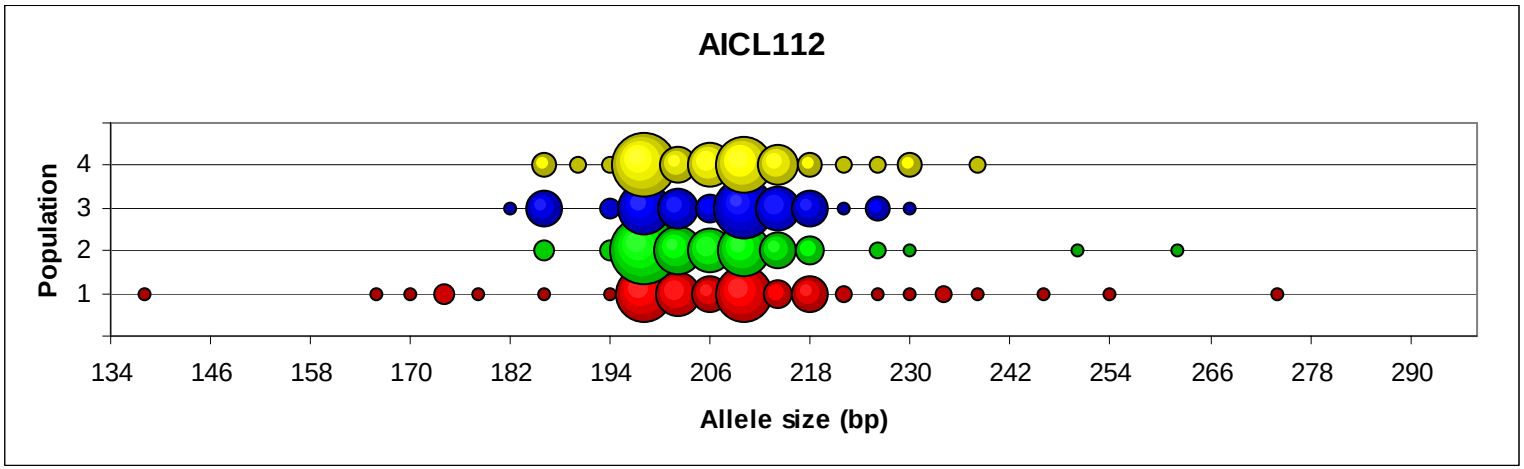


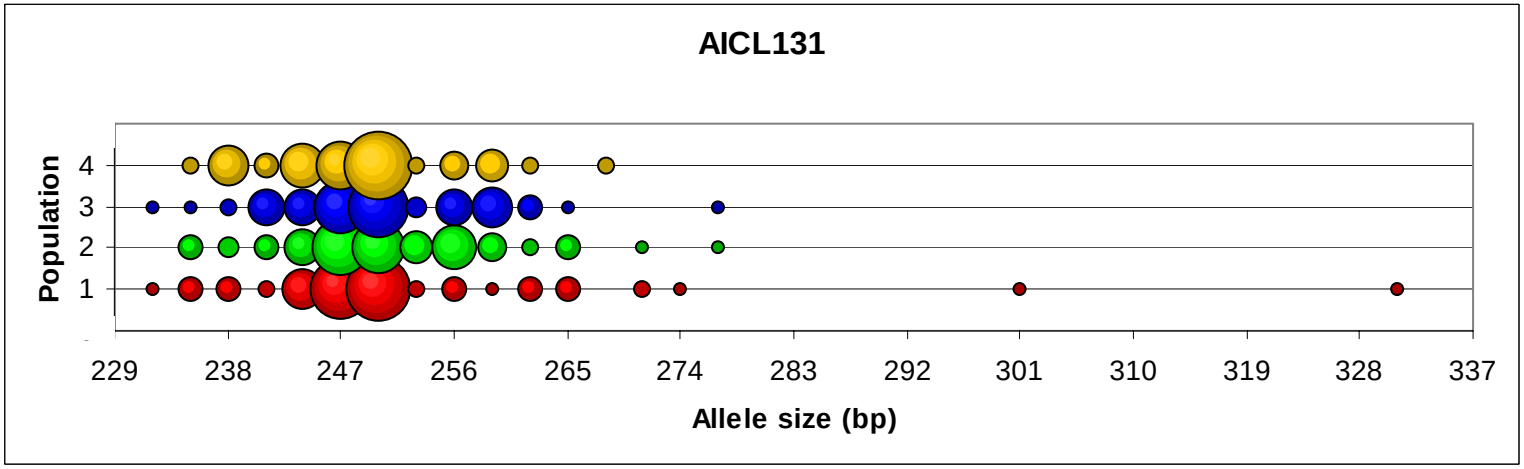

M26
Population
0
1
2
3
4
5
117
120
123
126
129
132
135
138
141
144
147
Allele size (bp)


G340
Population
0
1
2
3
4
5
102
105
108
111
114
117
120
123
126
129
Allele size (bp)


S336
Population
0
1
2
3
4
5
119
122
125
128
131
134
Allele size (bp)


N391
Population
5
4
3
2
1
0
221
227
233
239
245
251
257
263
269
275
281
Allele size (bp)


AICL112
Population
4
3
2
1
134
146
158
170
182
194
206
218
230
242
254
266
278
290
Allele size (bp)


AICL131
Population
4
3
2
1
229
238
247
256
265
274
283
292
301
310
319
328
337
Allele size (bp)