Vol. 423: 131–142, 2011
doi: 10.3354/meps08922

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published February 10

Microsatellite loci indicate population structure and selection between Atlantic and Gulf of Mexico populations of the bay scallop *Argopecten irradians*

E. M. Hemond[1,2,*], A. E. Wilbur[1]

[1]Center for Marine Science, University of North Carolina Wilmington, Wilmington, North Carolina 28409, USA
[2]Marine Science Center, Northeastern University, Nahant, Massachusetts 01908, USA

ABSTRACT: Declines in populations of bay scallops over the last 3 decades have spurred interest in their conservation and restoration. The development of sound restoration and management strategies requires a clear understanding of the genetic diversity and structure of populations throughout the species' range. In an effort to assess the population genetic structure of bay scallops, we evaluated 6 populations in the western North Atlantic and the Gulf of Mexico using 9 microsatellite markers, 5 of which are presented here for the first time. Results of traditional *F*-statistics indicate strong population structure between scallops in the Gulf of Mexico and western North Atlantic ($F_{ST} \geq 0.120$) and a low level of differentiation between North Atlantic (New York and North Carolina, USA) populations ($F_{ST} = 0.035$). However, distinct differences in high-frequency alleles at a single locus were largely responsible for the dramatic population genetic structure observed between Gulf and Atlantic populations. Further evaluation of loci by comparing observed data with simulated neutral distributions of F_{ST} indicates that selection is likely acting on the variation at 3 loci, while the variation at the other 6 loci does not deviate from the expectations of neutrality. Results suggest that the genetic structure between populations of bay scallops in the Gulf and Atlantic is affected by both selective pressure and restricted gene flow and that caution should be taken to maintain the native genetic diversity during commercial or restoration-based aquacultural activities.

KEY WORDS: Bay scallop · *Argopecten irradians* · Microsatellites · Population structure · Selection

INTRODUCTION

Bay scallops *Argopecten irradians* 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, decreases in the abundance of scallops throughout this range have been observed. This has been attributed to a number of factors including overfishing, loss of seagrass habitat, diminished water quality, and microalgal blooms (Tettelbach & Wenczel 1993, Arnold 2001).

Argopecten irradians is an epibenthic bivalve inhabiting sheltered inshore habitats (<12 m depth), particularly seagrass beds (Thayer & Stuart 1974). Bay scallops are not sessile like many bivalve species, but the distance over which they travel as adults is limited. They are simultaneous hermaphrodites with a broadcast spawning reproductive strategy, and dispersal and gene flow between populations most likely occur during a 1 to 2 wk planktonic larval stage (Castagna & Duggan 1971), which provides the potential for substantial genetic exchange between populations. The patchy distribution of this species throughout its range may be the result of barriers to dispersal or the lack of suitable habitat in some geographical regions.

The presence of marked differentiation in morphology among *Argopecten irradians* from different parts of the geographical range has led to the recognition of 3 subspecies. These include *A. i. irradians* (Lamarck,

*Email: hemond.e@neu.edu

1819), inhabiting the Atlantic coast from New Hampshire to New Jersey, *A. i. concentricus* (Say, 1822), inhabiting the coast from New Jersey south, including the Outer Banks of North Carolina to Florida and the eastern coastal Gulf of Mexico, and *A. i. amplicostatus* (Dall, 1898), ranging west of Louisiana to Galveston, Texas, and possibly farther south (Fig. 1). A fourth subspecies specific to the Florida Keys, *A. i. taylorae* (Petuch, 1987), was described based on morphological differences, but its status has been questioned due to the results of genetic analysis and a re-evaluation of morphology (Blake & Graves 1995, Marelli et al. 1997a,b). The most notable distinctions between subspecies are variation in plicae (ribs), convexity of the left valve, and pigmentation of the right valve. These and other general shape characteristics have been shown to be heritable (Kraeuter et al. 1984, Adamkewicz & Castagna 1988, Wilbur & Gaffney 1997), which suggests some degree of genetic differentiation supporting these subspecies classifications.

The subspecific categorization of *Argopecten irradians* has been complicated by phylogeographic studies using genetic markers that have revealed contradictory information. While 16S restriction fragment length polymorphism (RFLP) analysis supported the current classification of *A. i. concentricus* in North Carolina and Florida, and *A. i. irradians* as a divergent group (Wilbur 1995), a different study of mtDNA whole genome RFLPs (Blake & Graves 1995) showed less differentiation between Atlantic samples of *A. i. concentricus* and *A. i. irradians* than between Atlantic and Gulf *A. i. concentricus* populations (for a review, see Beaumont 2000). A possible genetic break between Atlantic and Gulf of Mexico populations of this estuarine species would follow a pattern observed in numerous marine and terrestrial species with ranges in the southeastern United States (Avise 1992) but would contradict the current subspecific classification of *A. irradians*.

In this study, we investigated the genetic diversity and population structure of this species throughout its range using 9 nuclear microsatellites, 5 of which are reported for the first time here. Microsatellites have become the genetic marker of choice for studies of population genetic structure because they are biparentally inherited, have a high mutation rate, and are presumably selectively neutral, although recent research has called this into question (Nielsen et al. 2006, Larsson et al. 2007). We used the 9 microsatellite markers, traditional *F*-statistics, and Bayesian clustering analysis to compare 3 populations of *Argopecten irradians* from Florida in the Gulf of Mexico and 3 populations from the Atlantic (2 in North Carolina and 1 in New York). We predicted that they would show 1 of 3 possible patterns of genetic diversity: (1) range-wide genetic homogeneity, (2) greater differences between than within subspecies (i.e. Florida and North Carolina scallops are more similar than either is to New York scallops), or (3) divergence between Gulf of Mexico and Atlantic populations (i.e. New York and North Carolina scallops are more similar to each other than to Florida populations).

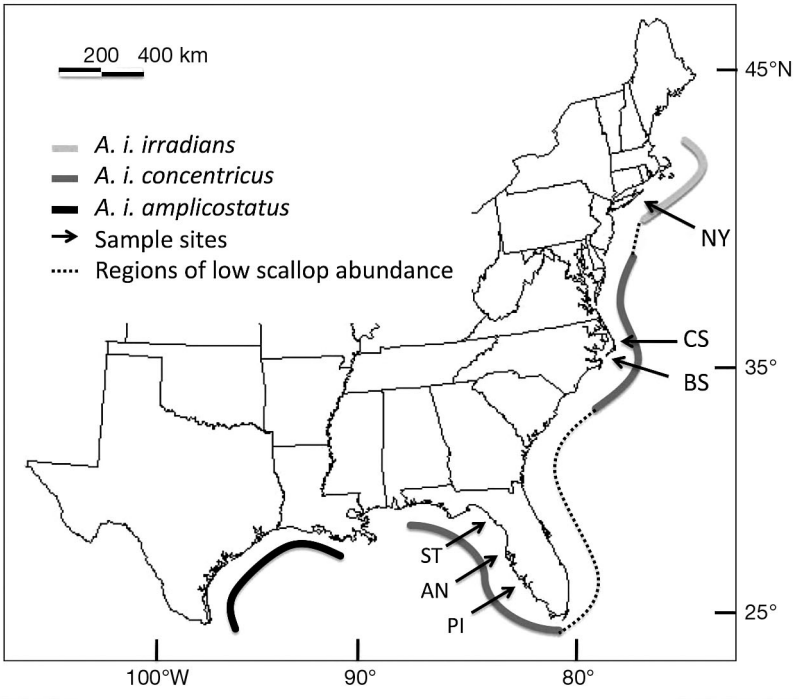

Fig. 1. *Argopecten irradians.* Scallop collection sites, with ranges of subspecies. Gulf of Mexico collection sites include Steinhatchee (ST), Anclote Estuary (AN), and Pine Island Sound (PI), Florida. Atlantic collection sites include Bogue Sound (BS) and Core Sound (CS), North Carolina, and Long Island, New York (NY)

Population genetic structure, as determined by the differentiation of allelic frequencies between populations, can result from genetic drift acting on neutral alleles or from selection on fitness-related alleles. Historically, presumed neutral molecular markers (particularly allozymes, mitochondrial DNA, or microsatellites) have been used to infer levels of gene flow among populations, because they should reflect population demography and phylogenetic history more accurately than selected loci (Luikart et al. 2003). However, studies have revealed that some of these markers may, in fact, not be neutral (Karl & Avise 1992, Ruiz-Pesini et al. 2004,

Nielsen et al. 2006), and, additionally, that regions of the genome subject to selection may be more useful than neutral ones for determining biologically relevant population structure and conservation units (Vasemägi & Primmer 2005, André et al. 2010). Under neutral theory, most allele frequencies in a population will shift randomly by genetic drift and will diverge over time in populations not connected by gene flow. However, when alleles are under selective pressure, either balancing selection, which maintains high allelic diversity, or directional selection, in which specific alleles are favored due to a fitness advantage, allele frequencies in different populations may be more or less differentiated than they would be due to genetic drift. This can result in local adaptation of populations as well as dramatic changes in the frequencies of alleles linked to those under selective pressure (Beaumont 2005). In studies with the goal of understanding the population dynamics of a species, it is important to test loci for the influence of selection, as population genetic structure based on selection may have different evolutionary and conservation implications than that based on genetic drift (Luikart et al. 2003). In addition to investigating patterns of genetic diversity, we applied a genome scan analysis to our microsatellite data to detect loci showing a signature of selection.

MATERIALS AND METHODS

Data collection. Genotyping of scallops was conducted using 9 microsatellite loci, including 5 loci (AICL112, AICL115, AICL131, AICL271, and AICL327; Table 1) that we developed using a microsatellite enriched DNA library (Glenn & Schable 2005) obtained with the assistance of the Savannah River Ecological Laboratory, and 4 loci (M26, G340, S336, and N391) from Roberts et al. (2005). The primers for the latter loci were modified from their published sequence by the addition of a 5′ pig-tail to promote adenylation and reduce the presence of stutter peaks (Brownstein et al. 1996). Genotypes were obtained for 3 population samples collected from the Gulf coast of Florida (Steinhatchee in 1998, N = 25; Anclote Estuary in 2001, N = 50; Pine Island Sound in 2001, N = 50), 2 population samples from North Carolina (collected in 1998 from Bogue Sound, N = 24; and Core Sound, N = 25), and 1 population sample from New York (collected in 1999 from Sag Harbor, Long Island, N = 47; Fig. 1). Adductor muscle was removed and stored at −80°C until extraction. DNA was extracted using a Puregene DNA Purification Kit (Gentra Systems), and PCRs were carried out in 25 µl reactions containing 1× PCR buffer, 1.5 mM $MgCl_2$, 2 µM dNTPs (1.2 µM for AICL112 and AICL327), 0.4 µM of each primer (0.2 µM for AICL115 and AICL327), 0.75 to 1 U *Taq*, and 50 to 100 ng template DNA. Thermal PCR conditions were as follows: 5 min at 94°C, followed by 40 cycles of 30 s at 94°C, 30 s at primer-specific annealing temperature (Table 1), 60 s at 72°C, and a final extension step at 72°C for 5 min. After confirmation of amplification on agarose gels, PCR products were diluted 1:10 to 1:100 with H_2O (according to the intensity of the band when electrophoresed through a 2% agarose gel stained with ethidium bromide), and 1 µl diluted PCR product was added to 9 µl Hi-Di:ROX solution (with a ratio of 1000:25) and visualized on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Resulting peaks were analyzed using GeneScan 3.7 and Genotyper 3.7 software (Applied Biosystems). Due to either stutter or presence of an insertion/deletion (indel) causing intermediate peaks in data for locus N391 (a di-nucleotide simple sequence repeat, SSR), we binned alleles as a tetra-nucleotide repeat; this did not affect Hardy-

Table 1. Summary information for microsatellite loci. Bracketed primer basepairs identify pig-tails (Brownstein et al. 1996) added to minimize stutter and enhance resolution of genotypes. Size range indicates the observed allele size range in the data. Repeat indicates simple sequence repeat (SSR) motifs observed in the sequenced clone from which the primers were designed. F: forward; R: reverse

Locus	Primers	Size range (bp)	Annealing temp. (°C)	Repeat	GenBank accession no.
AICL112	F: TGCCAAATCCATTTGCATATTA R: [GT]TTCCCTGTTCACTTGACAGACC	138–294	56	GACA	HQ593146
AICL115	F: TGCGGTATTTGAGTCCCCTA R: [GT]TTGACCTTTTGACCCCAAAT	157–225	56	GTCT	HQ593147
AICL131	F: CCCTATGGCTTCCTCAACCT R: [GT]TTAACTTTCTGTGCCGTGGA	232–331	50	CAA	HQ593148
AICL271	F: CCTTACATGACCCTGGCTGT R: [GT]TTCATCTAATTTATCAACCGACCA	71–107	50	CAAA	HQ593149
AICL327	F: GCAAAATCCACCCATCAGTT R: [GTTT]ACCGGAGGGGACTAGTGTTT	86–114	58	CAGA	HQ593150

Weinberg equilibrium. All data were scanned with MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) to identify any potentially misrecorded alleles and to look for signs of null alleles, stutter peaks, and large allele dropout, which could have led to genotyping errors.

Data analysis. Collected data were analyzed using the Genetic Analysis in Excel 6 (GenAlEx) add-in (Peakall & Smouse 2006) to determine allelic frequencies in each population. Number of alleles, allelic richness adjusted for sample size (El Mousadik & Petit 1996), inbreeding coefficient (F_{IS}, Wright 1951), and probability of Hardy-Weinberg disequilibrium (heterozygote deficiency) were calculated in FSTAT 2.9.3 (Goudet 1995). Linkage disequilibrium was assessed for each population in GENEPOP 4.0 (Raymond & Rousset 1995). Hardy-Weinberg disequilibrium and linkage disequilibrium were assessed against an alpha-level of 0.05 adjusted with a sequential Bonferroni procedure (Rice 1989). For linkage disequilibrium, the correction was conducted over all tests within each population (number of tests simultaneously carried out = $k = 36$, smallest adjusted $\alpha/k = 0.0014$); for Hardy-Weinberg disequilibrium, the correction was conducted by locus ($k = 6$, smallest adjusted $\alpha/k = 0.0083$).

Pairwise F_{ST} (Infinite Alleles Model, Wright 1951) between all populations was calculated for all loci and for putative selectively neutral loci using FSTAT, and significance was determined with 15 000 simulations. Based on the absence of significant population genetic structure among Florida populations or between North Carolina populations, these samples were collapsed into a regional Florida sample and a regional North Carolina sample for subsequent analyses. Global and pairwise estimators of population differentiation F_{ST} (θ, Weir & Cockerham 1984) and R_{ST} (Stepwise Mutation Model, Slatkin 1995) (ρ, Rousset 1996 as estimated by Michalakis & Excoffier 1996) were calculated in GENEPOP for each locus individually and all loci together for 3 regional populations (Florida, North Carolina, and New York).

The application LOSITAN (Antao et al. 2008) was employed to run the FDIST 2 program (Beaumont & Nichols 1996), which identifies loci under selection using the F_{ST} outlier method. This method uses observed data to generate an expected distribution of F_{ST} versus heterozygosity and identifies loci that fall outside of a chosen confidence interval as candidates for selection. This program is based upon the Lewontin-Krakauer test (Lewontin & Krakauer 1973, and see Beaumont 2005), formulated on the idea that genetic structure among populations should be reflected to a similar degree by all selectively neutral loci, whereas loci under directional selection (or linked to loci under selection) should exhibit significantly greater differentiation, and those loci under balancing selection should exhibit less differentiation than expected. Pairwise tests were conducted for both the infinite alleles model (IAM, Kimura & Crow 1964) and the stepwise mutation model (SMM, Kimura & Otha 1978) using 50 000 realizations, the 'neutral mean F_{ST}' option, and both the 99% and 95% confidence intervals (CI). Using the neutral F_{ST} approximation, any locus identified as an outlier is removed and the test is then rerun. Presence of outlying loci in the analysis causes a bias in the mean and variance of F_{ST} values resulting in a failure to identify less extreme outliers as candidates for selection. By removing the extreme outliers and recalculating mean F_{ST} values, additional candidate loci can be identified.

Bayesian clustering of individuals was conducted using STRUCTURE 2.0 (Pritchard et al. 2000). This method groups individuals into the number of clusters selected by the user so as to minimize linkage disequilibrium and Hardy-Weinberg disequilibrium within each cluster. STRUCTURE then calculates a posterior probability, $P(X|K)$, of the observed genotype data (X) given the number of clusters (K) and calculates a membership coefficient (Q) for each individual to each of the constructed clusters. We used a burn-in period of 50 000 with a run time of 100 000, an admixture model, putative population information not included, and correlated allele frequencies. The program was run for a range of clusters (K) from 1 to 6, and 5 replicates were run for each value of K to verify consistency of the results. The posterior probability values were used to determine the number of clusters that most accurately fit and reflect the data. The program was run for 3 different sets of data based upon the results of the neutrality tests: (1) all loci, (2) 'selected' loci (those with significantly greater than average F_{ST} values) between Atlantic and Gulf of Mexico populations (AICL327 and N391), and (3) all 'selectively neutral' loci (those loci with no evidence of selection acting on their alleles: AICL112, AICL115, AICL131, AICL271, G340, and S336).

RESULTS

Population structure

Following sequential Bonferroni corrections, a significant heterozygote deficiency was observed for 1 locus, AICL115, in Steinhatchee (Table 2). No cases of linkage disequilibrium were significant. No significant genetic differentiation was observed among the 3 Florida populations (Table 3) or between North Carolina populations (all loci $F_{ST} = 0.004$). All other pairwise comparisons among populations showed significant differences, regardless of whether all loci or only

Table 2. *Argopecten irradians.* Statistics of microsatellite loci for the number of individuals scored (N), the number of alleles observed (N_a), allelic richness (R), F_{IS}, and p-values for heterozygote deficiency in population samplings. For population source abbreviations, see Fig. 1. *p < 0.05 after sequential Bonferroni correction

Locus	Source	N	N_a	R	F_{IS}	p
AICL112	AN	50	21	13.76	0.129	0.016
	PI	50	12	9.74	0.047	0.267
	ST	25	13	12.59	0.072	0.243
	BS	24	14	13.79	−0.056	0.908
	CS	25	13	12.82	−0.01	0.671
	NY	47	14	12.11	0.094	0.056
AICL115	AN	50	12	8.67	0.162	0.028
	PI	49	10	8.47	0.072	0.204
	ST	25	7	6.76	0.366	0.003*
	BS	23	7	7.00	0.041	0.459
	CS	25	6	5.84	−0.046	0.720
	NY	47	9	7.08	0.002	0.561
AICL131	AN	49	16	12.17	0.161	0.009
	PI	50	13	11.22	0.036	0.311
	ST	25	11	10.68	−0.095	0.940
	BS	24	8	7.83	−0.01	0.627
	CS	25	9	8.60	0.171	0.087
	NY	47	8	6.08	0.02	0.473
AICL271	AN	50	9	8.31	−0.056	0.859
	PI	50	9	7.73	−0.042	0.795
	ST	25	7	6.76	−0.1	0.879
	BS	24	8	7.88	−0.02	0.668
	CS	25	6	5.92	−0.118	0.919
	NY	47	6	5.81	0.044	0.360
M26	AN	50	5	4.42	−0.191	0.990
	PI	50	7	5.73	−0.089	0.883
	ST	25	6	5.92	−0.061	0.794
	BS	24	5	4.96	0.078	0.357
	CS	25	4	4.00	0.123	0.254
	NY	47	6	4.98	−0.125	0.951
G340	AN	50	7	5.90	0.137	0.072
	PI	50	5	4.98	−0.102	0.929
	ST	25	7	6.92	−0.125	0.950
	BS	24	5	5.00	−0.27	0.999
	CS	25	6	5.84	0.001	0.583
	NY	47	5	4.71	−0.153	0.964
N391	AN	50	10	9.49	0.008	0.504
	PI	47	10	9.57	0.085	0.091
	ST	25	8	8.00	0.183	0.025
	BS	24	8	7.92	−0.051	0.768
	CS	25	9	8.75	−0.045	0.769
	NY	47	8	6.95	−0.003	0.583
S336	AN	48	3	2.73	0.129	0.212
	PI	49	3	2.47	0.206	0.088
	ST	25	3	2.92	−0.111	0.801
	BS	24	4	3.96	−0.145	0.896
	CS	24	3	3.00	0.217	0.127
	NY	47	3	3.00	−0.07	0.796
AICL327	AN	50	6	4.75	0.019	0.509
	PI	50	6	4.72	−0.024	0.710
	ST	25	5	4.83	−0.063	1.000
	BS	24	3	2.96	0.332	0.061
	CS	25	3	2.92	0.053	0.520
	NY	47	6	4.95	−0.044	0.731

'selectively neutral' loci were included in the analysis (Table 3). All samples from Florida and all samples from North Carolina were collapsed into 2 regional populations for additional analyses. Florida, North Carolina, and New York regional populations were used to evaluate population genetic structure between the Gulf of Mexico and the Atlantic.

F_{ST} and R_{ST} estimates of regional population differentiation across all loci indicated differences among the 3 regions. However, there was a discrepancy between these 2 measures of differentiation (Table 4). Based on estimates of F_{ST}, North Carolina and New York scallops (F_{ST} = 0.035) were less differentiated than either was relative to scallops from the Gulf of Mexico ($F_{ST} \geq 0.120$). Results of R_{ST}, however, indicated a similar level of divergence between Florida and New York scallops (0.061) and North Carolina and New York scallops (0.057), while divergence between Florida and North Carolina scallops was much higher (0.158). F_{ST} is based on the IAM (Kimura & Crow 1964), while R_{ST} is based on the SMM (Kimura & Otha 1978), which is expected to reflect the mutation pattern of microsatellites. Under the SMM, alleles that are closer in size share a closer ancestry than those alleles separated by a large size difference.

The discrepancy between R_{ST} and F_{ST} can be understood by looking at the allele frequencies at locus AICL327 (Fig. 2). North Carolina and New York shared the same 2 high-frequency alleles, 86 basepairs (bp) and 102 bp, while Florida was dominated by the high frequency of the 98 bp allele. As a result, F_{ST} for the comparison of New York versus North Carolina was much smaller than for either Atlantic population relative to Florida (Table 3). However, R_{ST} was smaller between New York and Florida scallops because of the frequency differences of the 86 and 102 bp alleles in New York and North Carolina scallops and because of the smaller length difference between the 98 and 102 bp alleles, compared to that of the 86 bp allele. Because the vari-

Table 3. *Argopecten irradians.* Pairwise F_{ST} between all sampled populations. F_{ST} calculated with all loci are shown above the diagonal and those calculated with 'selectively neutral' loci are below the diagonal. Significance was calculated with 15 000 simulations. For population source abbreviations, see Fig. 1.
$^{*}p < 0.01$, $^{**}p < 0.001$

		FL			NC		NY
		AN	PI	ST	BS	CS	
FL	AN		0.002	0.001	0.127**	0.116**	0.114**
	PI	0.003		−0.001	0.146**	0.133**	0.123**
	ST	−0.001	−0.005		0.147**	0.138**	0.125**
NC	BS	0.044**	0.078**	0.066**		0.004	0.042**
	CS	0.024**	0.052**	0.043**	0.012		0.031**
	NY	0.028**	0.052**	0.045**	0.029**	0.006*	

ance is often higher for R_{ST} than for F_{ST} (Balloux & Lugon-Moulin 2002) and the results of Bayesian clustering conform more closely to the results of F_{ST} than R_{ST}, we do not consider R_{ST} further in the interpretation of our results.

Measures of pairwise F_{ST} calculated by locus can be understood by looking at the variation in allele frequencies (Table 4, Fig. 2). F_{ST} values indicated similarities between New York and North Carolina scallops for 5 loci (AICL112, AICL271, AICL327, N391, and S336), similarity between Florida and New York scallops for 3 loci (AICL115, AICL131, G340), and similarity between Florida and North Carolina scallops for 1 locus (M26). Locus AICL327 revealed the greatest population structure, largely due to the near-fixation of the 98 bp allele in Florida scallops, which was absent in North Carolina scallops and represented by only a single occurrence in New York scallops. Furthermore, in both the North Carolina and New York populations, the same 2 alleles, 86 and 102 bp, were most frequent but occurred at low frequencies (0.07 and 0.01, respectively) in the Florida population. Despite the relative similarity between North Carolina and New York scallops at locus AICL327, the F_{ST} value (0.169) was still highly significant.

Allele frequencies at other loci provide further evidence of population structure. The distribution of allele frequencies at locus N391 was also distinct, as the 249 bp allele occurred at similarly high frequencies in North Carolina (0.48) and New York (0.52) scallops and at a frequency of only 0.10 in Florida scallops. Looking at locus AICL112, the 198 bp allele was the most frequent allele in Florida scallops (frequency of 0.22), but was almost absent in North Carolina scallops (frequency of 0.02) and was not found in New York scallops. Further similarities occur in the distribution of low-frequency, larger alleles in the North Carolina and New York populations, in contrast to frequencies observed in the Florida population.

Tests of neutrality

Results for the IAM and SMM runs were comparable. In all pairwise population comparisons, the F_{ST} values for AICL327 were dramatically higher than the F_{ST} values for all other loci (Table 4). Tests of neutrality implicate this locus as a candidate for selection in all pairwise comparisons among regional populations, falling outside the 99% CI for selectively neutral loci in all 3 cases (Fig. 3). In addition, locus N391 fell outside the 99% CI for the Florida/New York comparison (Fig. 3A) and the 95% CI for the Florida/North Carolina comparison (Fig. 3B). In the New York/North Carolina comparison, locus M26 revealed marginal significance, falling outside the 95% CI (Fig. 3C). The other 6 loci appeared to fall within the range of neutral differentiation expected for selectively neutral loci based on genetic drift.

Table 4. *Argopecten irradians.* F_{ST} (θ) and R_{ST} (ρ) coefficients between Florida (FL), North Carolina (NC), and New York (NY) and over all populations for each individual locus and all loci. Significance of F_{ST} based on exact *G*-tests of genetic differentiation.
$^{*}p < 0.01$, $^{**}p < 0.001$

Locus	FL/NC		FL/NY		NC/NY		Global	
	F_{ST}	R_{ST}	F_{ST}	R_{ST}	F_{ST}	R_{ST}	F_{ST}	R_{ST}
AICL112	0.037**	0.280	0.045**	0.150	−0.003	0.034	0.034**	0.197
AICL115	0.048**	−0.004	0.004	−0.003	0.015	−0.012	0.025**	−0.005
AICL131	0.027**	−0.006	0.025**	0.018	0.018	0.012	0.025**	0.006
AICL271	0.008**	0.004	0.026**	0.002	−0.001	−0.013	0.014**	0.001
AICL327	0.655**	0.508	0.576**	0.013	0.169**	0.255	0.557**	0.306
M26	0.024*	0.052	0.055**	0.003	0.081**	0.088	0.046**	0.037
G340	0.069**	0.042	0.023**	−0.006	0.030	0.032	0.045**	0.021
N391	0.099**	0.056	0.115**	0.048	−0.006	−0.009	0.093**	0.046
S336	0.124**	0.034	0.127**	0.033	0.035*	0.086	0.111**	0.045
All loci	0.132**	0.158	0.120**	0.061	0.035**	0.057	0.112**	0.106

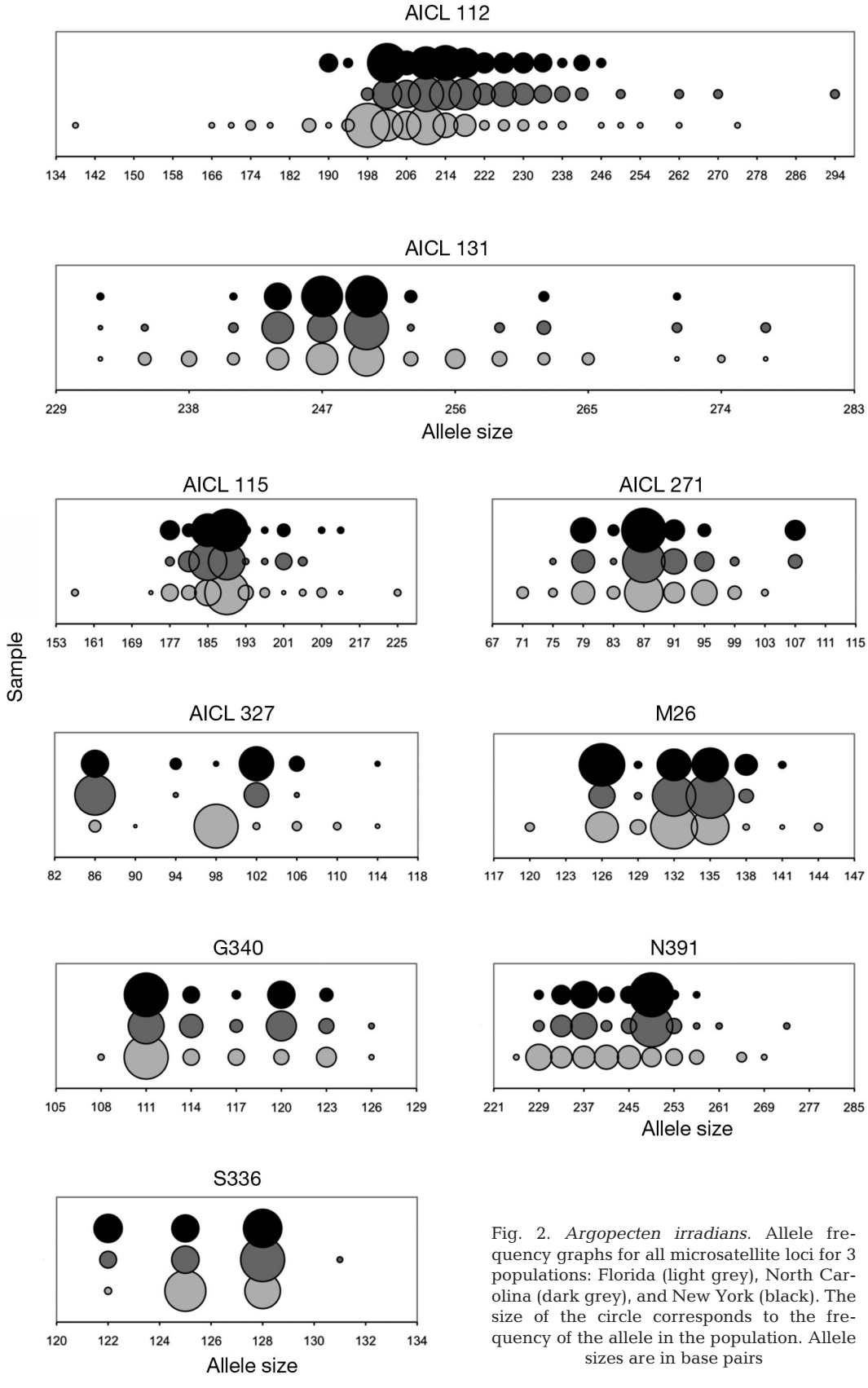

Fig. 2. *Argopecten irradians*. Allele frequency graphs for all microsatellite loci for 3 populations: Florida (light grey), North Carolina (dark grey), and New York (black). The size of the circle corresponds to the frequency of the allele in the population. Allele sizes are in base pairs

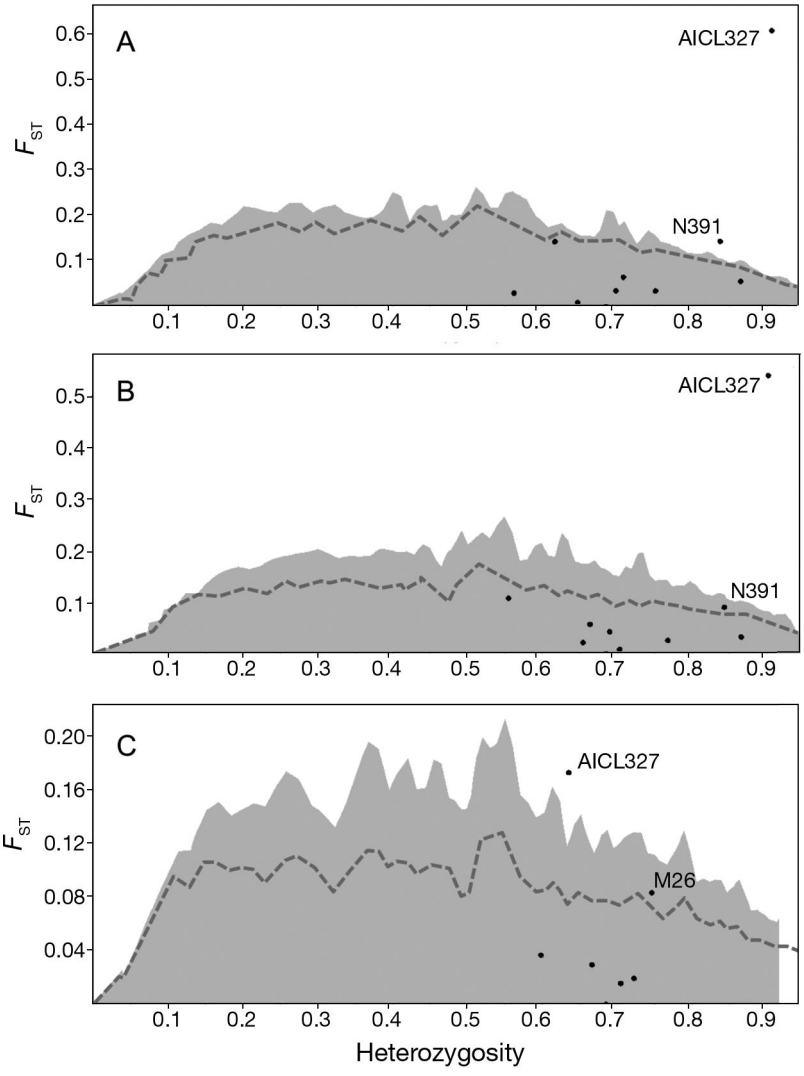

Fig. 3. *Argopecten irradians.* Results for neutrality tests on all 9 loci for 3 pairwise comparisons between (A) Florida and New York, (B) Florida and North Carolina, and (C) New York and North Carolina. Loci within the shaded region demonstrate F_{ST} vs. heterozygosity values expected due to neutral variation; loci above the shaded region (99% CI) or dashed line (95% CI) are candidates for selection

Bayesian cluster analysis

The results of the neutrality tests were used to group loci for the 3 different analyses in STRUCTURE (see 'Materials and methods'). Analysis based on all loci revealed 2 distinct clusters dividing the Gulf and Atlantic regions: one for individuals originating in Florida and a second for those originating in New York and North Carolina (Fig. 4A). Classification of the data as 2 groups ($K = 2$) resulted in high probability of membership values for each population to either Gulf or Atlantic cluster (mean $Q > 0.95$), indicating a strong ability to discriminate between individuals from these 2 regions. The log likelihood values were similar for $K = 2$ and $K = 3$; however, these data analyzed for 3 groups greatly decreased the probability of membership (Q) values. For interpretation of the data, we followed the recommendation of Pritchard & Wen (2004) in choosing 'the smallest value of K that captures the majority of the data.' When the 2 loci (AICL327 and N391), which are candidates for selection between the Gulf and Atlantic, were applied to the analysis, results were very similar, with high ability to discriminate between these regions (Fig. 4B, mean $Q > 0.94$). The final scenario using only 'selectively neutral' loci (without AICL327, N931, and M26) revealed low ability to discriminate between scallops originating in Atlantic or Gulf regions (mean $Q = 0.50$ to 0.77 when $K = 2$; Fig. 4C). For this final analysis, the posterior probability for the observed data set was higher for $K = 1$ and $K = 3$ than for $K = 2$, but probability of membership values remained low (mean $Q = 0.69$ to 0.72 when $K = 3$; data not shown).

DISCUSSION

Similarity among populations within regions

For many marine species, a pelagic larval stage provides the potential to travel over great distances and homogenize the gene pools of geographically distinct populations; however, recent studies have shown or predicted that local recruitment often dominates over long-distance dispersal (Cowen et al. 2000, Taylor & Hellberg 2003). Based on recruitment patterns, Arnold et al. (1998) found that populations of *Argopecten irradians* along the Gulf coast of Florida, from Anclote to Steinhatchee, are unlikely to share a large number of migrants. Similarly, 2 studies (Peterson & Summerson 1992, Peterson et al. 1996) found predominantly localized recruitment of *A. irradians* in North Carolina. Isolation of North Carolina populations in Bogue and Back Sounds was corroborated by mtDNA evidence showing significant population genetic structure (AMOVA index $\Phi_{CT} = 0.0145$) between those basins (Marko & Barr 2007). Our analyses do not support isolation of populations within regions, as our 9 micro-

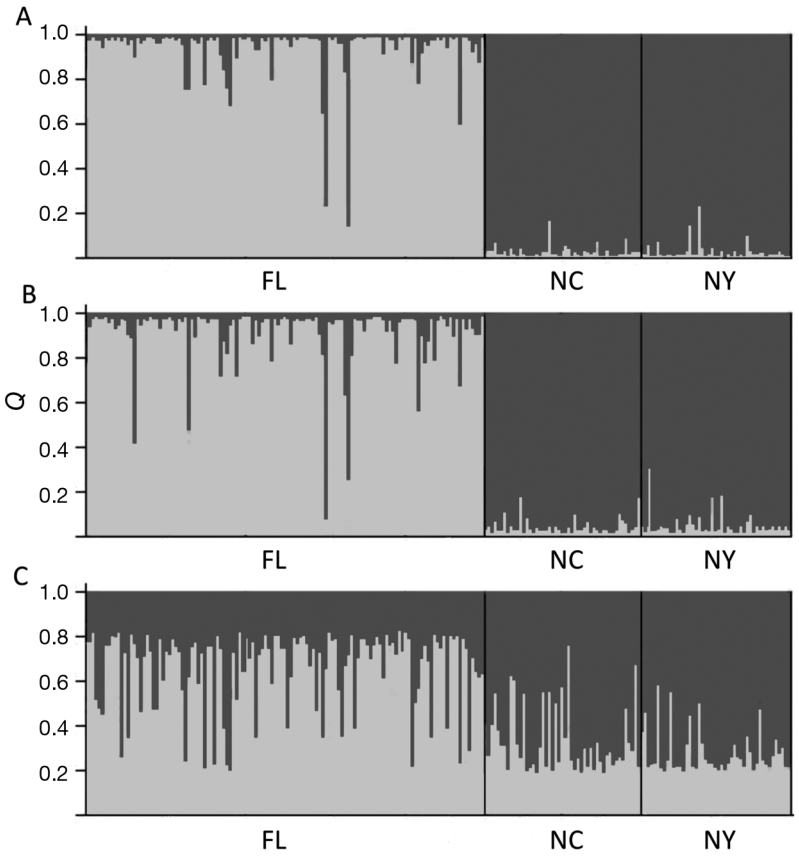

Fig. 4. *Argopecten irradians.* Probability of membership (*Q*) values for individual scallops from Florida (FL), North Carolina (NC), and New York (NY) populations to 2 clusters (Cluster 1: light grey, Cluster 2: dark grey) using STRUCTURE. Results represent membership values when clusters were constructed using (A) all loci, (B) candidate selected loci (AICL327 and N391), and (C) 'selectively neutral' loci

satellites indicated little to no population genetic structure among the 3 *A. irradians* populations in the Gulf of Mexico or between the 2 North Carolina populations sampled. It is likely that the 2-fold lower effective population size of mtDNA relative to nuclear DNA in hermaphrodites may be responsible for this discrepancy, since the rate of genetic drift acting on allele frequencies should be greater for mtDNA than for nDNA. However, a direct comparison using microsatellites to evaluate scallops between Bogue and Back Sounds would be necessary to understand the dynamics of these 2 populations, which showed significant differentiation at the mtDNA locus.

Population genetic structure between the Gulf of Mexico and the Atlantic

Highly significant genetic differences were observed in each of the comparisons between samples in the Gulf of Mexico and Atlantic, as well as significant, but less dramatic genetic differentiation between North Carolina and New York. This finding supports the mtDNA data of Blake & Graves (1995) and allows us to reject the hypothesis of range-wide genetic homogeneity in this species. Furthermore, our data do not support the current subspecies classification of *Argopecten irradians.* While morphological data group Florida and North Carolina populations together as *A. i. concentricus* and populations from Maryland and New Jersey to Cape Cod as *A. i. irradians* (Beaumont 2000, Blake & Shumway 2006), microsatellite data indicate greater genetic divergence between the 2 sampled populations of *A. i. concentricus* than between the 2 Atlantic populations.

The observation of a genetic break in microsatellite variation between Gulf of Mexico and Atlantic populations is consistent with a pattern of divergence in the mitochondrial lineages of numerous species (Avise 1992), including marine and coastal terrestrial animals with very different life history characteristics. The similarity of these genetic breaks along the east coast of Florida could reflect either current restricted gene flow between populations north and south of this boundary or historical vicariance with relatively recent secondary contact. The latter explanation is related to the geographic history of the southeastern US, 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. Contemporary isolation of Gulf of Mexico and Atlantic bay scallops may be maintained by the absence of preferred seagrass habitat along the Atlantic coast of Florida, or by the ocean circulation patterns in the Gulf of Mexico and Florida Straights, particularly eddies that can entrain larvae and prevent migration. However, the extreme differences in the magnitude of structure among loci in the present study suggest the presence of additional factors influencing differentiation, such as selection.

Genetic drift versus selection

Given geographic isolation of populations, frequencies of neutral alleles should shift within populations as a result of genetic drift. Due to the randomness of genetic drift and variation in mutation rates among loci, some alleles will be affected more rapidly by drift than others. A representative sampling of loci will provide a distribution of the variation in neutral genetic divergence between populations. Loci falling far outside of this distribution, in terms of F_{ST} versus heterozygosity, are candidates for being under selective pressure or genetic hitchhiking with a locus under selection (Lewontin & Krakauer 1973, Maynard Smith & Haigh 1974, Beaumont & Nichols 1996). While microsatellites have been considered selectively neutral markers for population genetic studies, recent examples have revealed that this is often not the case. Selection and selective sweeps have been implicated in inflating F_{ST} estimates in Atlantic cod (Nielsen et al. 2006) and herring (Larsson et al. 2007). Three of the microsatellite loci evaluated here were identified as potential markers of selective differentiation: 2 loci (AICL327 and N391) between the Atlantic and the Gulf of Mexico, and 2 loci (AICL327 and M26) between New York and North Carolina.

While this is not the first study to use a limited number of microsatellite markers to identify selection (Nielsen et al. 2006, Larsson et al. 2007), and it is important in studies of population genetic structure to test for the influence of selection, we remain cautious in concluding that the 2 marginally outlying loci (N391 and M26) are candidates for selection with using only 9 markers as a reference. A dataset including 10-fold or 100-fold as many loci would provide a much more robust neutral distribution of F_{ST} versus heterozygosity, but this is prohibitively expensive for the current study with this non-model organism. However, the F_{ST} values across all pairwise comparisons for locus AICL 327 were so high that the locus would remain an outlier unless the distribution of F_{ST} values for the other loci shifted dramatically. Thus, this locus is the strongest candidate for selection.

While AICL327 and N391 indicated a potential role of divergent selection in the genetic differentiation between Gulf and Atlantic populations, all of the 'selectively neutral' loci also indicated significant population genetic structure. Selection may be driving the dramatic genetic differences between these regions, but significant genetic differentiation due to genetic drift still reflects a barrier to gene flow between Gulf and Atlantic populations. Despite the frequently observed genetic break between the Gulf of Mexico and Atlantic for many species, this has not previously been attributed to divergent selection. A genomic scan of eastern oyster *Crassostrea virginica* population genetic structure between the Atlantic and Gulf of Mexico using amplified fragment length polymorphisms (AFLPs) found 3 loci potentially under selective pressure (Murray & Hare 2006); however, these did not remain significant candidates for selection after correcting for multiple tests.

Population structure between Atlantic populations

In addition to notable genetic distinctions between Gulf of Mexico and Atlantic bay scallop populations, 3 loci (AICL327, M26, and S336) suggest population genetic differentiation between New York and North Carolina scallops. However, Bayesian clustering did not clearly distinguish between these 2 Atlantic samples, and F_{ST} values for most loci support the little genetic separation between New York and North Carolina. Of the 3 loci that do indicate structure between these 2 populations, 2 are potentially influenced by selection.

Recent work by Campanella et al. (2007) used 8 microsatellite loci (including 4 of the same markers used in our study: M26, G340, N391, and S336) to evaluate the origin of a recently repopulated scallop population in New Jersey. They were able to distinguish between the 2 potential source populations, North Carolina and Long Island, New York, due to private alleles at loci S336, M26, and C1832 (not included in our study) in North Carolina. More genetic similarity was observed between New Jersey and New York than North Carolina, in contrast to the results of an mtDNA RFLP study conducted for the same geographic areas in 1998 (Bologna et al. 2001). It has therefore been proposed that stocks of bay scallops in New Jersey are being naturally supplemented with larval transport from Long Island populations, and that this represents a shift from recruitment from North Carolina populations. However, at least one of these microsatellite loci (M26) may be affected by selection or genetic hitchhiking, potentially making it a non-neutral marker for evaluating gene flow. Currents may facilitate larval exchange between North Carolina and New York populations, perhaps through intermediate populations, while selective pressure due to differences at these locations may maintain differences in allele frequencies at certain loci. Further evaluation of these markers with additional loci will be necessary to determine whether they are acting neutrally and can be used to draw conclusions regarding recruitment.

Implications and further research

Discrepancy between patterns observed for different loci in this species indicate the need to use a diversity

of genetic markers in studies of population structure, as well as the importance of testing the loci for signatures of selection. Further analysis of genetic structure among these populations using additional microsatellite markers and over a number of years will provide a more accurate indication of the distribution of F_{ST} at neutral markers and the temporal stability of allele frequencies. In addition, a more continuous sampling of populations at intermediate geographic locations would indicate whether the genetic break between Atlantic and Gulf populations occurs over a localized area, or over a cline of shifting allele frequencies, although, due to the absence of scallops along the southeastern US, the potential for such an analysis is limited.

The microsatellite loci designed by Roberts et al. (2005) and used in this study were obtained from an expressed sequence tag (EST) database, and are therefore in coding regions of DNA. Those designed by us for the purpose of this study were obtained through cloning of a microsatellite-enriched genomic DNA library, which can include loci in either coding or non-coding regions of DNA. We conducted a BLAST search in the NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine if any of the 5 novel loci used for this study could be matched to any known sequence. It is significant that the clone sequence used to design primers for locus AICL327 was found to be similar to an EST sequence (ID: CK484191.1; Expectation (E-value) = 1×10^{-90}), identified as bay scallop veliger larvae mRNA. This sequence did not include the microsatellite repeat and did not match any known proteins in the NCBI database. Because AICL327 includes a tetra-nucleotide repeat, it is unlikely to be in an expressed region itself, but it is possible that it is 'hitchhiking' with a closely linked locus under selection. Investigation of the functionality of this associated EST gene, as well as genes in proximity to N391 and M26, is a possibility for further study.

With regard to current aquacultural activities involving *Argopecten irradians*, our results provide substantial evidence of strong allele frequency shifts along the range of this species, implicating either restricted gene flow between populations or selection acting on expressed genes, or both. The genetic structure observed between populations thought to represent a single subspecies indicates that caution should be taken in performing any transfer of animals between regions. Particularly if selection is acting to remove alleles from the population, transplant activities may introduce less fit individuals into recipient populations.

Acknowledgements. We thank W. S. Arnold and S. Geiger (FFWRI) for providing bay scallop samples from Florida, P. Murphy (NCDMF) for scallops from North Carolina, and S. Tettelbach (LIU) for scallops from New York. We also appreciate the comments of R. Bilgin, D. Combosch, S. Vollmer and 6 anonymous reviewers on analyses and/or earlier versions of this manuscript. A portion of this work was supported by the NMSF Saltonstall-Kennedy Project grant no. Na17FD2368.

LITERATURE CITED

Adamkewicz SL, Castagna M (1988) Genetics of shell color in the bay scallop, *Argopecten irradians*. J Hered 79:14–17

André C, Larsson LC, Laikre L, Bekkevold D and others (2010) Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity, doi: 10.1038/hdy.2010.71

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. BMC Bioinformatics 9:323

Arnold WS (2001) Bivalve enhancement and restoration strategies in Florida, U.S.A. Hydrobiologia 465:7–19

Arnold WS, Marelli DC, Bray CP, Harrison MM (1998) Recruitment of bay scallops *Argopecten irradians* in Floridian Gulf of Mexico waters: scales of coherence. Mar Ecol Prog Ser 170:143–157

Avise JC (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, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. Mol Ecol 11:155–165

Beaumont A (2000) Genetic considerations in transfers and introductions of scallops. Aquacult Int 8:493–512

Beaumont MA (2005) Adaptation and speciation: What can F_{ST} tell us? Trends Ecol Evol 20:435–440

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc Biol Sci 263:1619–1626

Blake SG, Graves JE (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 NJ, Shumway SE (2006) Bay scallop and calico scallop fisheries, culture and enhancement in eastern North America. In: Shumway SE, Parsons GJ (eds) Scallops: biology, ecology and aquaculture, 2nd edn. Developments in aquaculture and fisheries science, Vol 35. Elsevier B.V., Amsterdam, p 945–964

Bologna PAX, Wilbur AE, Able KW (2001) Reproduction, population structure, and recruitment limitation in a bay scallop (*Argopecten irradians* Lamarck) population from New Jersey, USA. J Shellfish Res 20:89–96

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004–1006, 1008–1010

Campanella JJ, Bologna PAX, Kim LEJ, Smalley JV (2007) Molecular genetic evidence suggests Long Island as the geographic origin for the present population of bay scallops in Barnegat Bay, New Jersey. J Shellfish Res 26: 303–306

Castagna M, Duggan W (1971) Rearing the bay scallop, *Aequipecten irradians*. Proc Natl Shellfish Assoc 61:80–85

Cowen RK, Lwiza KMM, Sponaugle S, Paris CB, Olson DB (2000) Connectivity of marine populations: open or closed? Science 287:857–859

El Mousadik A, Petit RJ (1996) Chloroplast DNA phylogeography of the argan tree of Morocco. Mol Ecol 5:547–555

Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. Methods Enzymol 395:202–222

Goudet J (1995) FSTAT version 1.2: a computer program to calculate *F*-statistics. J Hered 86:485–486

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. Science 256: 100–102

Kimura M, Crow JF (1964) The number of alleles that can be maintained in finite populations. Genetics 49:725–738

Kimura M, Otha T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. Proc Natl Acad Sci USA 75:2868–2872

Kraeuter J, Adamkewicz SL, Castagna M, Wall R, Karney R (1984) Rib number and shell color in hybridized subspecies of the Atlantic bay scallop, *Argopecten irradians.* Nautilus 98:17–20

Larsson LC, Laikre L, Palm S, Andre C, Carvalho GR, Ryman N (2007) Concordance of allozyme and microsatellite differentiation in a marine fish, but evidence of selection at a microsatellite locus. Mol Ecol 16:1135–1147

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74:175–195

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4:981–994

Marelli DC, Krause MK, Arnold WS, Lyons WG (1997a) Systematic relationships among Florida populations of *Argopecten irradians* (Lamark 1819) (Bivalvia: Pectinidae). Nautilus 110:31–41

Marelli DC, Krause MK, Lyons WG, Arnold WS (1997b) Subspecific status of *Argopecten irradians concentricus* (Say, 1822) and the bay scallops of Florida. Nautilus 110:42–44

Marko PB, Barr KR (2007) Basin-scale patterns of mtDNA differentiation and gene flow in the bay scallop *Argopecten irradians concentricus.* Mar Ecol Prog Ser 349:139–150

Maynard Smith J, Haigh J (1974) Hitch-hiking effect of a favorable gene. Genet Res 23:23–35

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

Murray MC, Hare MP (2006) A genomic scan for divergent selection in a secondary contact zone between Atlantic and Gulf of Mexico oysters, *Crassostrea virginica*. Mol Ecol 15:4229–4242

Nielsen EE, Hansen MM, Meldrup D (2006) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. Mol Ecol 15:3219–3229

Peakall R, Smouse PE (2006) GenAlEx 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

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

Pritchard JK, Wen W (2004) Documentation for *structure* software: Version 2. University of Chicago, Chicago, IL. Available at http://pritch.bsd.uchicago.edu.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Rice WR (1989) Analyzing tables of statistical tests. Evolution 43:223–225

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

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142:1357–1362

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004) Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303:223–226

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Taylor MS, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. Science 299:107–109

Tettelbach ST, Wenczel P (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 12:423–431

Thayer GW, Stuart HH (1974) The bay scallop makes its bed of seagrass. Mar Fish Rev 35:27–30

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Vasemägi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Mol Ecol 14:3623–3642

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370

Wilbur AE (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, DE

Wilbur AE, Gaffney PM (1997) A genetic basis for geographic variation in shell morphology in the bay scallop, *Argopecten irradians.* Mar Biol 128:97–105

Wright S (1951) The genetical structure of populations. Ann Eugen 15:323–353

Editorial responsibility: Inna Sokolova, Charlotte, North Carolina, USA

Submitted: June 30, 2010; Accepted: November 5, 2010
Proofs received from author(s): January 21, 2011