

Contemporary and historical influences on the genetic structure of the estuarine-dependent Gulf killifish *Fundulus grandis*

Dean A. Williams^{1,2,*}, Stacy D. Brown¹, Douglas L. Crawford¹

¹Marine Biology and Fisheries, Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, Florida 33149, USA

²Present address: Department of Biology, 2800 S. University Dr., Texas Christian University, Fort Worth, Texas 76129, USA

ABSTRACT: In comparison to species living in open marine environments, estuarine-dependent species are expected to exhibit stronger genetic population structure due to dispersal limitations. Estuarine habitats are relatively transitory on geological time scales; thus, populations may not be at migration–drift equilibrium, which could confound estimates of current day gene flow or selection. We used 8 nuclear microsatellite loci to investigate the genetic structure of the estuarine Gulf killifish *Fundulus grandis* across 10 populations along the northwestern and northeastern Gulf of Mexico. Patterns of isolation by distance, spatial autocorrelation, and assignment tests indicate that dispersal is limited and occurs primarily between neighboring sites. Principal component analysis and Bayesian clustering revealed evidence for genetic discontinuities located in Mobile Bay and western Florida which are near hypothesized biogeographical boundaries. There was also a significant negative relationship between genetic diversity and latitude, a pattern consistent with the presence of hypothesized refugia in the southern Gulf regions during the Pleistocene that later recolonized the northern Gulf. Results suggest that populations may be at or near migration–drift equilibrium at a regional scale (e.g. the western Gulf), but that dispersal barriers and potential historical signatures on population structure will need to be taken into consideration at larger spatial scales.

KEY WORDS: Genetic structure · Isolation by distance · Equilibrium · Estuary · Microsatellites · Gulf of Mexico

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Marine fish inhabiting the open ocean often exhibit low genetic population differentiation due to their large effective population sizes and high dispersal potential (Ward et al. 1994, Waples 1998, Grosberg & Cunningham 2001). In contrast, species that utilize near-shore habitats such as estuaries and lagoons are expected to exhibit more genetic structure than strictly marine species due to the discontinuous nature of these habitats and their relative isolation from ocean currents (Bilton et al. 2002, Watts & Johnson 2004). Decreased levels of genetic exchange combined with variable environmental conditions such as salinity and

temperature should enhance the genetic divergence and local adaptation of populations in estuarine areas (e.g. Lemaire et al. 2000, Beheregaray & Sunnucks 2001).

At migration–drift equilibrium, dispersal between these near-shore habitats should approximate a model in which individuals are most likely to disperse to neighboring populations along a single dimension, thereby producing a genetic pattern of isolation by distance (IBD) (Kimura & Weiss 1964, Crow & Aoki 1984, Gold & Turner 2002, Burridge et al. 2004). The transitory distribution of these estuarine habitats on geological time scales, however, may prevent these fish populations from attaining migration–drift equilibrium

*Email: dean.williams@tcu.edu

and prevent the formation of IBD genetic structure (Chenoweth & Hughes 2003, Pampoulie et al. 2004, Durand et al. 2005). Determining the scale over which migration–drift equilibrium may exist in these systems is important since equilibrium is assumed when estimating migration rates using estimators of population subdivision such as Wright's fixation index (F_{ST}), or when inferring a signal of natural selection relative to neutral expectations (e.g. Whitehead & Crawford 2006, Duvernell et al. 2008).

The Gulf of Mexico's coastline and its connections to the Atlantic have undergone many changes. During the Pliocene (>1.8 million years before present [mybp]), sea levels were much higher, much of the Florida landmass was underwater, and the connection between the Atlantic Ocean and Gulf of Mexico was more expansive. During the Pleistocene (1 800 000 to 11 550 ybp), glaciation captured and stored much water and sea level dropped up to 150 m (McIntyre et al. 1976, Bloom 1983, Imbrie et al. 1983). With lower sea levels, Florida's shoreline extended 100 miles (ca. 161 km) west of its current position, which effectively isolated the Gulf of Mexico's marine populations from Atlantic populations, resulting in a strong genetic discontinuity for many taxa between the Gulf and the Atlantic (Reeb & Avise 1990, Avise 2000, Soltis et al. 2006). Additionally, Florida's landmass was also much more arid during the glacial maxima (~18 000 ybp), which may have resulted in fewer and more isolated estuarine habitats along the Gulf shore (Avise 1996). During interglacial periods in the Pleistocene, sea levels rose, Florida was a series of islands, and the Gulf and Atlantic Ocean were once again highly connected. Some genetic discontinuities have been described in the northern Gulf (e.g. Felder & Staton 1994, Gurgel et al. 2004, Bilodeau et al. 2005). During the glacial maxima, temperatures along the northern continental margins of the Gulf coast were 4 to 5°C lower than they are today (Brunner 1982). January temperatures were as much as 17°C lower than at present in the southeastern United States at this time, suggesting that shallow coastal environments would also have been considerably colder (Stanley 1986). These low temperatures are hypothesized to have forced some near-shore taxa into southern refugia in Mexico and south Florida (Barnwell & Thurman 1984, Felder & Staton 1994). With the rise of temperature and sea levels during the interglacial periods, these populations would have expanded back into the northern Gulf. These changes in the Gulf's shoreline, temperature, and connections to the Atlantic should have affected the distribution and population demographics of species inhabiting estuaries.

The Gulf killifish *Fundulus grandis* is a common inhabitant of salt marshes along the coast from the St.

Johns River in northeastern Florida (30° 23' 55.93' N, 81° 23' 53.47' W) to Laguna de Tamiahua, Veracruz, Mexico (21° 35' 36.51' N, 97° 33' 07.73' W) (Lee et al. 1980). We used nuclear microsatellite loci to investigate the population genetic structure of *F. grandis* throughout much of its range in the Gulf of Mexico. *F. grandis* deposits its eggs in dense mats of marsh vegetation that are usually only flooded during biweekly high tides every 2 wk (Greeley & MacGregor 1983). After an incubation period of about 2 wk, the returning tide induces hatching and the larvae then complete their development in the intertidal marsh. This egg stranding reproductive strategy is predicted to promote the retention of offspring in the natal estuary and contribute to significant genetic structure between estuaries. Direct estimates of adult dispersal distance are not known for *F. grandis*. Mark-recapture dispersal estimates for the closely related *F. heteroclitus*, which inhabits salt marshes along the Atlantic coast of North America, ranged from 200 m to 3.6 km within a single year (Sweeney et al. 1998, Skinner et al. 2005). If the genetic structure of *F. grandis* is predominantly shaped by current-day dispersal and distributional patterns, we would expect to find a pattern of IBD across the Gulf with significant genetic differentiation among populations due to low levels of dispersal. Alternatively, if past climatic events have had a significant impact on the genetic structure of *F. grandis* in the Gulf, then we might expect to find a lack of IBD across the Gulf, evidence for secondary contact in the northern Gulf between previously separated populations, and evidence of a negative relationship between latitude and genetic diversity.

MATERIALS AND METHODS

Sampling. *Fundulus grandis* were collected from 10 locations across the Gulf of Mexico from August 2005 to October 2006 either using baited minnow traps and seine nets, or they were purchased from local bait shops (Fig. 1). We attempted to capture this species at several sites along the southwestern coast of Florida (near Naples) and on the eastern coast near Miami; however, we were not successful in capturing any individuals in these areas. Four samples (100% of Kiln County, 100% of Lake Charles, 30% of Dauphin Island, and 50% of Sarasota individuals; see Table 1) were obtained at bait shops. The owners could verify the capture site for all of these populations. A total of 440 individuals were analyzed, with 40 ± 5.47 (mean \pm SE) individuals collected from each location (see Table 1).

DNA was extracted from dried fin clips. The tissue was placed in 300 μ l lysis buffer (75 mM NaCl, 25 mM

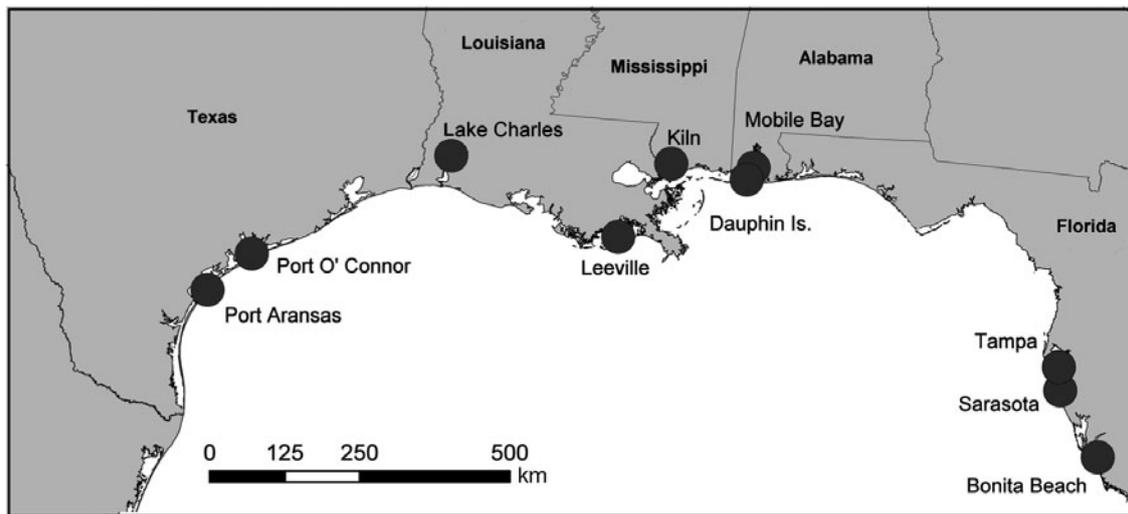


Fig. 1. Sampling localities of *Fundulus grandis* in the Gulf of Mexico

EDTA, 1% sodium dodecyl sulphate [SDS]) and incubated with 0.1 mg Proteinase K at 55°C for 2 h. Proteins were precipitated by adding a half volume of 7.5 M ammonium acetate followed by centrifugation for 10 min. DNA was precipitated from the supernatant by adding 0.7 volume of isopropanol followed by centrifugation for 15 min. The DNA pellet was washed with 70% ethanol, allowed to air dry for 30 min, and then resuspended in 50 µl 10 mM Tris-HCl pH 8.5.

DNA amplification and genotyping. We optimized 8 microsatellite loci previously characterized for *Fundulus heteroclitus* (Adams et al. 2005) for *F. grandis* in 3 fluorescently labeled multiplex primer groups containing the following final concentrations: A (0.05 µM ATG-B101, 0.05 µM ATG-25, 0.075 µM ATG-20), B (0.075 µM CA-1, 0.05 µM ATG-12), and C (0.06 µM ATG-18, 0.06 µM ATG-B4, 0.06 µM ATG-B103). Each 10 µl reaction contained 2.5 mM MgCl₂, 1X PCR buffer, 0.2 mM dNTPs, 0.4 U Taq DNA polymerase, 70 ng DNA, and one of 3 primer multiplexes (see above for concentrations). The PCR thermal cycling profile consisted of 94°C for 2 min, followed by 31 cycles of 94°C for 15 s, 55°C for 15 s, and 72°C for 30 s, ending with a 5 min extension step at 72°C. Following PCR amplification, the products were electrophoresed on an ABI 3730XL Genetic Analyzer (Applied Biosystems). All loci that did not amplify were reamplified in single locus reactions at least 2 times. Individuals that did not produce PCR products for more than half of the loci were discarded from the analyses.

We used GENEMAPPER version 4.0 (Applied Biosystems) to score the genotypes. All genotypes were checked by 2 individuals (D. A. Williams and S. D. Brown). To estimate our genotyping error rate, we extracted DNA from a duplicate set of fin tissue taken from a random subset of individuals ($n = 53$) across

most populations and genotyped them at all loci. All of the test genotypes were identical to previously determined genotypes, indicating that our error rate was very low (<2%).

Statistical analysis. We tested the loci for deviation from Hardy-Weinberg equilibrium and linkage disequilibrium using GENEPOP version 3.3 (Raymond & Rousset 1995). The number of alleles (N_A), observed heterozygosity (H_O), and expected heterozygosity (H_E) were calculated using GenAlEx6 (Peakall & Smouse 2006). Allelic richness (A_R) for each locality was calculated using FSTAT version 2.9.3 (Goudet 2001) with a sample size adjustment of $n = 17$ individuals (the smallest sample size).

We used the program MICRO-CHECKER (Van Oosterhout et al. 2004) to infer whether heterozygote deficits within a population may have been due to technical difficulties such as null alleles, scoring difficulties due to stutter, or large allele dropout. Alternatively, biological processes such as population substructuring (i.e. Wahlund effect), inbreeding, and genetic bottlenecks can also cause heterozygote deficits in populations. These processes should generally produce consistent heterozygote deficits across all or most loci, whereas technical difficulties usually only affect one or a few loci. When null alleles were inferred at a locus, we estimated their frequency using Brookfield's equation, which assumes that non-amplifications are a result of null allele homozygotes (Brookfield 1996). We then corrected all allele frequencies for that locus before estimating levels of genetic diversity and calculating levels of differentiation.

The program BOTTLENECK (Cornuet & Luikart 1996, Piry et al. 1999) tested for the genetic signature of a recent reduction in the effective population size (N_e) in all samples. Populations that have experienced

a reduction in N_e are expected to have excess H_E relative to that expected under mutation–drift equilibrium. This occurs because allelic richness is lost at a significantly faster rate than heterozygosity after a population reduction. We evaluated all 3 mutation models (infinite allele model [IAM], 2 phase mutation model [TPM], and stepwise mutation model [SMM]). A Wilcoxon sign-rank test was then used to determine if a significant number of loci exhibited excess heterozygosity (Cornuet & Luikart 1996). We also used a graphical method to look for evidence of a bottleneck by plotting the number of alleles in 7 allele frequency categories (Luikart et al. 1998). A population that has not experienced a bottleneck is expected to show an L-shaped distribution (many low frequency alleles and few high frequency alleles), whereas bottlenecked populations will exhibit a mode shift.

Differentiation among populations was calculated in FSTAT using the F_{ST} analog theta (Weir & Cockerham 1984). We tested for pairwise differences between populations using the log likelihood statistic G (Goudet et al. 1996) (not assuming Hardy-Weinberg equilibrium) in FSTAT. To illustrate the relationships among populations we used principal component analysis (PCA) of allele frequencies using PCAGEN (www2.unil.ch/popgen/softwares/pcagen.htm). The significance of the axes was calculated based on 10 000 permutations. We tested for IBD by correlating pairwise $F_{ST}/(1 - F_{ST})$ between populations with geographic distance using 1000 permutations in a Mantel test in GenAlEx6. We also tested for genetic spatial autocorrelation using GenAlEx6 (Peakall & Smouse 2006). The autocorrelation coefficient, r , is a measure of the genetic similarity between populations that fall within a defined distance class. The significance of r is determined by random permutation of all populations among distance classes and recalculating r 1000 times to set the upper and lower 95% confidence limits around this value. If the r value fell above or below these limits, then significant spatial structure was inferred. We also calculated 95% CI around each r value by bootstrapping r values within each distance class 1000 times. If these 95% CI did not include $r = 0$, then significant spatial structure was inferred. We also used a simple assignment test to ask whether more individuals would be assigned to their population of capture or to neighboring populations than to populations further away as expected by a stepping stone model of dispersal. Individuals were assigned to populations in which they had the highest log likelihood score using the Bayesian based assignment method of Rannala & Mountain (1997) in GeneClass2 (Cornuet et al. 1999, Piry et al. 2004).

A Bayesian clustering method implemented in the program STRUCTURE version 2.1 was used to estimate the number of populations (K) in the Gulf of

Mexico and to assign individuals to these populations (Pritchard et al. 2000, Pritchard & Wen 2003). The membership of each individual in a population was estimated using the ancestry coefficient q_i , which varies on a scale from 0 to 1.0, with 1.0 indicating full membership. We ran the Monte Carlo Markov Chain (MCMC) for 10^6 iterations following a burn-in period of 10^5 iterations for $K = 1$ to 10 using the correlated allele frequencies model and assuming admixture (the default values) 10 times. The most likely K was then estimated using the method of Evanno et al. (2005).

RESULTS

All loci were highly polymorphic, with an average of 22.1 ± 3.4 alleles per locus (range = 13 to 35 alleles) for all populations combined (Table 1). None of the pairwise comparisons between loci within populations exhibited significant genotypic linkage disequilibrium after Bonferroni correction ($p < 0.0001$). This suggests that the loci we used provide independent measures of genetic diversity and genetic structure. We found no significant heterozygote excess in any sample. However, there was a significant heterozygote deficit at locus *ATG-25* in 2 populations and at loci *ATG-B101*, *ATG-B4*, and *CA-1* in the Bonita Beach population (Table 1). The lack of a clear pattern across loci or populations suggests that genetic bottlenecks or population admixture were unlikely explanations for the observed heterozygote deficits. The MICROCHECKER analysis indicated that the most likely technical cause of these deficits was null alleles rather than problems associated with scoring stutter peaks or large allele drop out.

None of the tests for a genetic bottleneck was significant. There was not a significant excess of heterozygosity relative to equilibrium expectations under any mutation model nor was there evidence of a mode shift in the frequency distribution of alleles (data not presented).

Average allelic richness across the 8 loci decreased with increasing latitude ($y = -0.36x + 20.5$, $r^2 = 0.54$, $p = 0.01$) (Fig. 2a). The pattern of allelic richness with longitude revealed that the samples from southern Texas and Florida have the highest allelic richness and that allelic richness decreases in populations further north (Fig. 2b). This resulted in a U-shaped distribution of allelic richness with longitude in the Gulf and is best explained with a second order polynomial relationship ($y = 0.03x^2 - 4.5x + 207.4$, $r^2 = 0.87$, $p = 0.001$) (Fig. 2b).

There was significant genetic structure across all 10 populations ($F_{ST} = 0.041 \pm 0.02$ SE; 95% CI = 0.02–0.08). The average pairwise F_{ST} between populations was 0.05 ± 0.004 ($n = 45$, range = 0.003 to 0.11)

Table 1. *Fundulus grandis*. Genetic diversity measures at microsatellite loci within 10 populations of *F. grandis*. N = number of individuals genotyped, N_A = number of alleles at a locus, H_O = observed heterozygosity (values in bold represent significant heterozygote deficits), H_E = expected heterozygosity, F_{IS} = inbreeding coefficient (calculated after correction for possible null alleles, see 'Materials and methods'), A_R = allelic richness corrected for sample size using 17 individuals

Sample (abbreviation)	Statistic	ATG-12	ATG-18	ATG-20	ATG-25	ATG-B101	ATG-B103	ATG-B4	CA-1	Mean
Pt. Aransas (AR)	N	34	34	34	34	34	34	34	34	34.0
	N_A	12	8	10	13	9	20	20	22	14.3
	H_O	1.00	0.79	0.68	0.82	0.88	0.91	0.91	0.91	0.86
	H_E	0.89	0.82	0.65	0.86	0.80	0.92	0.92	0.92	0.85
	F_{IS}	-0.104	0.074	-0.030	0.060	-0.036	0.027	0.018	0.045	0.007
	A_R	11.0	7.1	7.6	10.6	7.0	15.3	15.6	17.7	11.5
Pt. O'Connor (OC)	N	43	43	43	43	43	43	43	43	43.0
	N_A	14	7	9	11	10	20	18	19	13.5
	H_O	0.86	0.77	0.67	0.91	0.81	0.88	0.95	0.91	0.85
	H_E	0.87	0.81	0.62	0.87	0.78	0.93	0.92	0.91	0.84
	F_{IS}	0.013	0.035	-0.076	-0.036	-0.080	0.063	-0.020	0.015	-0.011
	A_R	11.1	6.3	7.3	9.0	7.6	15.5	14.4	14.5	10.7
Lake Charles (LC)	N	36	36	36	36	36	36	36	36	36.0
	N_A	10	9	9	10	8	17	18	16	12.1
	H_O	0.81	0.72	0.44	0.92	0.75	0.92	0.94	0.86	0.80
	H_E	0.79	0.80	0.49	0.84	0.80	0.93	0.90	0.88	0.80
	F_{IS}	-0.002	0.109	0.121	-0.073	0.037	0.027	-0.036	0.003	0.023
	A_R	8.5	7.1	7.0	8.8	6.8	14.8	13.4	12.8	9.9
Leeville (LV)	N	59	59	59	59	59	59	59	59	59.0
	N_A	10	7	8	10	8	21	19	18	12.6
	H_O	0.81	0.83	0.46	0.75	0.73	0.97	0.88	0.83	0.78
	H_E	0.81	0.77	0.49	0.85	0.79	0.94	0.93	0.84	0.80
	F_{IS}	0.033	-0.080	0.072	0.076	0.088	-0.026	0.066	0.035	0.033
	A_R	8.2	6.1	5.8	8.2	6.7	15.8	15.0	11.3	9.6
Kiln County (KC)	N	124	123	124	124	124	124	124	123	123.8
	N_A	11	13	9	9	9	27	28	20	15.8
	H_O	0.85	0.77	0.59	0.87	0.81	0.93	0.94	0.85	0.83
	H_E	0.81	0.80	0.59	0.83	0.81	0.94	0.92	0.83	0.82
	F_{IS}	-0.038	0.044	0.008	-0.046	-0.001	0.015	-0.02	-0.027	-0.01
	A_R	8.29	7.26	5.43	7.74	6.64	16.00	14.73	10.22	9.54
Dauphin Island (DI)	N	23	23	23	22	22	23	23	23	22.8
	N_A	6	8	7	8	8	17	16	11	10.1
	H_O	0.87	0.83	0.78	0.86	0.64	1.00	1.00	0.65	0.83
	H_E	0.69	0.80	0.76	0.81	0.62	0.92	0.91	0.68	0.77
	F_{IS}	-0.243	-0.006	-0.008	-0.030	-0.038	-0.065	-0.078	0.017	-0.056
	A_R	5.5	7.6	6.2	7.7	7.4	15.0	14.7	9.0	9.1
Mobile Bay (MB)	N	55	55	55	55	55	55	55	55	55.0
	N_A	10	8	8	9	8	20	17	19	12.4
	H_O	0.82	0.69	0.86	0.78	0.73	0.93	0.95	0.84	0.82
	H_E	0.77	0.75	0.81	0.81	0.74	0.93	0.91	0.83	0.82
	F_{IS}	-0.056	0.073	-0.051	0.038	0.053	0.017	-0.027	-0.018	0.004
	A_R	7.2	7.2	6.7	7.5	6.4	15.3	13.1	11.5	9.4
Tampa Bay (TB)	N	17	17	17	17	17	17	17	17	17.0
	N_A	12	6	4	8	11	12	15	18	10.8
	H_O	0.94	0.88	0.29	0.94	0.88	0.77	0.94	0.94	0.82
	H_E	0.89	0.78	0.27	0.78	0.84	0.87	0.88	0.93	0.78
	F_{IS}	-0.034	-0.093	-0.110	-0.106	-0.028	0.135	-0.056	0.077	-0.027
	A_R	12	6	4	8	11	12	15	18	10.8
Sarasota (SS)	N	31	31	31	31	31	31	31	31	31.0
	N_A	15	7	4	8	13	14	12	17	11.3
	H_O	0.94	0.81	0.68	0.61	0.97	0.77	0.71	0.94	0.80
	H_E	0.89	0.75	0.61	0.84	0.89	0.87	0.77	0.92	0.82
	F_{IS}	-0.039	-0.061	-0.081	0.117	-0.073	0.141	0.113	-0.032	0.011
	A_R	12.4	6.2	4.0	8.5	11.2	12.2	10.1	14.3	9.9
Bonita Beach (BB)	N	17	18	18	18	18	18	18	18	17.9
	N_A	12	6	5	10	8	15	12	18	10.8
	H_O	0.78	0.79	0.79	0.94	0.94	0.67	0.68	0.79	0.80
	H_E	0.90	0.69	0.67	0.86	0.84	0.90	0.86	0.94	0.83
	F_{IS}	0.168	-0.107	-0.144	-0.074	-0.101	0.174	0.246	0.192	0.044
	A_R	12.0	5.9	4.9	9.9	7.9	14.0	11.6	17.6	10.5

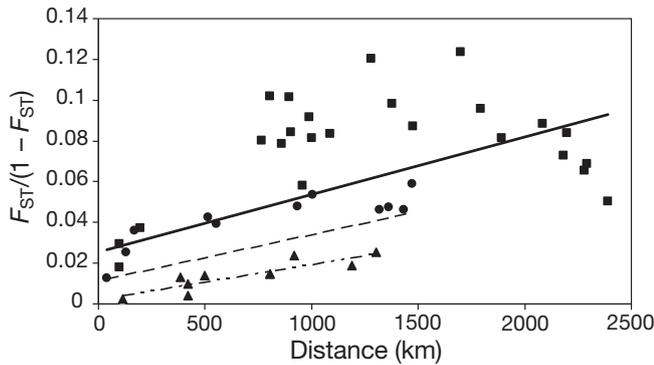


Fig. 4. *Fundulus grandis*. Relationship between pairwise genetic and geographic distance for 10 populations of *F. grandis*. Solid line is regression line for all 10 populations (\blacksquare , \bullet , \blacktriangle , $y = 3 \times 10^{-5}x + 0.025$); dashed line is 7 populations excluding the Florida populations (\bullet , \blacktriangle , $y = 2 \times 10^{-5}x + 0.011$); dot-dashed line is 5 populations west of Mobile Bay (\blacktriangle , $y = 2 \times 10^{-5}x + 0.002$). \blacksquare : pairwise comparisons between Florida and all other populations; \bullet : comparisons excluding the Florida populations; \blacktriangle : comparisons excluding Mobile Bay and Florida populations. F_{ST} : Wright's fixation index

There was a significant positive relationship between genetic and geographic distance ($r = 0.57$, $p = 0.01$) (Fig. 4). Most of the Florida comparisons with the western Gulf, however, form a cluster of points above the regression line, suggesting that these populations are more differentiated from the rest of the Gulf than predicted by distance alone. At the furthest distances, F_{ST} drops below the predicted relationship which may indicate the presence of homoplasy. There was still a significant relationship between genetic and geographic distance after removing the Florida samples ($r = 0.61$, $p = 0.02$), although these comparisons revealed that the Mobile Bay samples cluster above the predicted relationship (Fig. 4). Comparing only the populations west of Mobile Bay resulted in a strong relationship between genetic and geographic distance ($r = 0.89$, $p = 0.02$) with an origin very close to zero as expected when populations with a stepping stone population structure are at migration–drift equilibrium (Fig. 4). The effect of genetic differentiation between the Mobile Bay and Florida populations with the western Gulf populations can be seen in Fig. 4 by an increase in the y -intercept of the regression line for genetic versus geographic distance and is consistent with the groupings revealed by the PCA.

There was significant positive genetic spatial autocorrelation for pairwise F_{ST} values across the entire Gulf only in the 160 km ($p = 0.001$) and 480 km ($p = 0.003$) distance classes (Fig. 5), which generally encompasses 2 to 3 of our sampling sites. There was a steep decline in autocorrelation values for populations separated by distance classes of 726 km or more, with the correlogram crossing the y -axis at ~ 977 km and

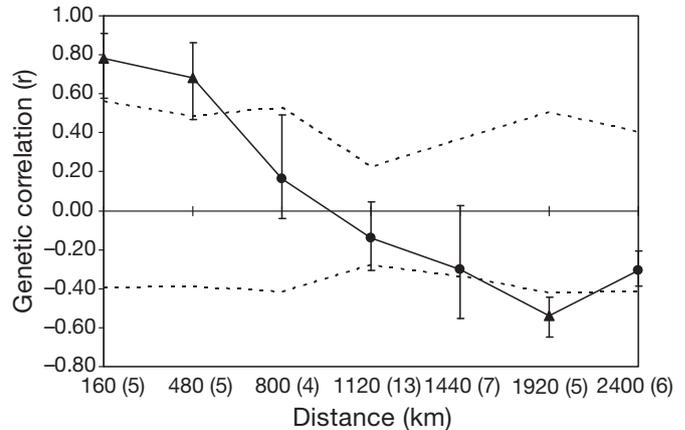


Fig. 5. *Fundulus grandis*. Genetic spatial autocorrelation of pairwise genetic (F_{ST}) and geographic distance for 10 populations of *F. grandis*. Dashed lines are 95% CI around $r = 0$; vertical lines are 95% bootstrapped CI around each calculated r value. \blacktriangle : significantly different from $r = 0$ ($p \leq 0.005$). Numbers in parentheses next to the distance classes are the number of pairwise comparisons in that class

becoming significantly negative at the 1920 km distance category, reflecting the distinct population clusters of Florida versus the western Gulf revealed by the PCA analysis and STRUCTURE (see below). Restricting the spatial autocorrelation only to the populations west of the Mobile Bay area revealed significant positive spatial autocorrelation in the first of 4 distance classes (160 km, $r = 0.68$, $p = 0.02$, data not presented).

One assumption of the individual-based assignment tests, such as the assignment method of Rannala & Mountain (1997) and the algorithm in STRUCTURE, is that the populations are in Hardy-Weinberg equilibrium. We therefore performed the analyses with and without locus *ATG-25*, which exhibited putative null alleles, since it is not possible to correct allele frequencies for null alleles and retain the correct individual genotypes. We also performed the analyses with and without the Bonita Beach population because it had heterozygote deficits at 3 loci. Excluding locus *ATG-25* or the Bonita Beach population had virtually no effect on the final results. The assignment test of Rannala & Mountain (1997) supported an IBD relationship. Most individuals were assigned to their population of capture (55.4%) or to an adjacent population (29.5%). STRUCTURE estimated 3 population clusters in the Gulf (Fig. 6). The Florida samples formed the strongest cluster, with an average proportion of ancestry (q) of 89.8% for Cluster A (Fig. 6). Although the other populations were a mix of the 3 clusters, the proportions of each cluster were strongly related to geography and exhibited a pattern consistent with IBD. Excluding the Florida populations (which were predominantly Cluster A), there was a strong positive relationship be-

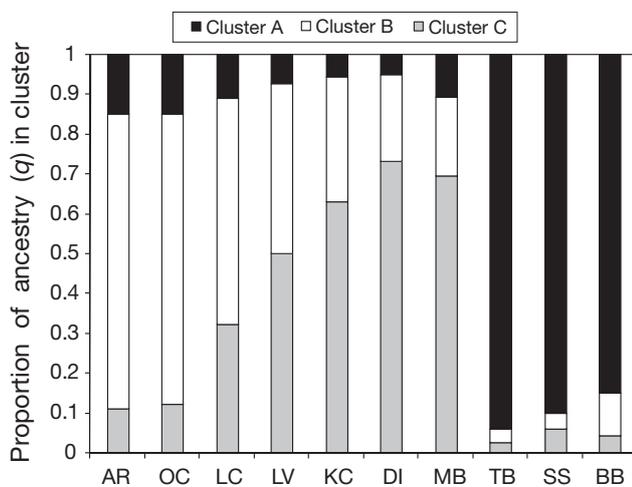


Fig. 6. *Fundulus grandis*. Proportion of ancestry (q) in each of the 3 clusters determined with STRUCTURE (see 'Material and methods') for each sampling locality of *F. grandis*. Abbreviations as in Table 1

tween longitude and the q that populations contained of Clusters A and B ($r = 0.87, 0.99, p < 0.01$, respectively) and a negative relationship between longitude and q that populations contained of Cluster C ($r = -0.99, p < 0.01$).

DISCUSSION

Fundulus grandis exhibits significant population genetic structure throughout the Gulf of Mexico. Our data suggest that most dispersal occurs between adjacent estuaries and that several genetic discontinuities exist along the Gulf coast. The Florida populations in particular are significantly differentiated from the more western and northern Gulf populations. Comparisons between the most geographically distant populations (Florida and the western Gulf) revealed a higher similarity in allele frequencies than predicted by distance, which may indicate the occurrence of homoplasy. When effective population sizes are very large, which may be the case for this species, homoplasy is predicted to occur in highly polymorphic markers such as microsatellites (Estoup et al. 2002). Populations at varying scales exhibited a pattern of IBD, and spatial autocorrelation revealed that populations located within 160 km of each other were more similar than expected at random. Similarly, assignment tests placed most individuals at either their place of capture or in adjacent populations. The clusters determined by the STRUCTURE analysis were also consistent with IBD in the western Gulf.

Estuarine-dependent species, like the Gulf killifish, are expected to exhibit stronger genetic population

structure due to dispersal limitations than species living in more open marine environments. Although limited in number, studies of estuarine fish in the Gulf of Mexico have found the expected IBD relationship (Gold & Richardson 1998, Blandon et al. 2001, Gold & Turner 2002, present study), whereas studies of reef or pelagic organisms in the Gulf have found broad-scale genetic homogeneity (Herke & Foltz 2002, Ball & Chapman 2003, Purcell et al. 2006).

Fundulus grandis shows very similar levels of population subdivision when compared to its closely related congener *F. heteroclitus* which inhabits salt marshes along the Atlantic coast of North America (mean pairwise $F_{ST} = 0.05 \pm 0.004$ for *F. grandis* vs. 0.042 ± 0.005 for southern populations of *F. heteroclitus*) (Adams et al. 2006, present study). *F. heteroclitus* exhibits concordant clines in allele frequencies at allozyme, mitochondrial, and nuclear microsatellite loci, with a strong transition zone occurring at $\sim 41^\circ$ N latitude (reviewed in Adams et al. 2006, Duvernell et al. 2008). Strong patterns of IBD across the entire range of *F. heteroclitus* ($r = 0.85$) as well as within northern ($r = 0.70$) and southern populations ($r = 0.79$) suggests that this species was distributed along the entire east coast of North America during the Pleistocene glacial maxima and that subsequent rising sea levels brought previously separated populations into contact and produced the observed transition zone (Adams et al. 2006).

Previous intraspecific studies on a number of taxa have found little evidence for genetic discontinuities in the Gulf of Mexico (but see Gurgel et al. 2004), although this may have resulted from a variety of factors including limited sampling and studies of species with high dispersal potential (e.g. pelagic fish) (Avisé et al. 1987, Reeb & Avisé 1990, Gold & Richardson 1998, Heist & Gold 2000, Blandon et al. 2001, Gold & Turner 2002, Herke & Foltz 2002, Ball & Chapman 2003). The 2 genetic discontinuities we found occur near previously described biogeographic boundaries (Mobile Bay and Florida). The presence of a vicariance zone has been described for some closely related species of fish and invertebrates near Mobile Bay in the northern Gulf (e.g. Bert 1986, McClure & McEachran 1992, Felder & Staton 1994, Gurgel et al. 2004, Harrison 2004, Bilodeau et al. 2005). The Florida sampling localities are along the northern edge of the Caribbean faunal biogeographic province, which is located just below the more temperate Gulf of Mexico province (Briggs 1974, Engle & Summers 2000).

We did not find a strong transition zone in the northern Gulf near Mobile Bay, and Bayesian clustering did not reveal a clear signal of admixture in this area arising from 2 differentiated populations coming into secondary contact as is observed in *Fundulus heteroclitus* (Adams et al. 2006). These results suggest that if *F.*

grandis had been forced into southern refugia (e.g. Barnwell & Thurman 1984), it may not have been for long periods of time or, alternatively, there may have been many episodes of extinction-recolonization in the northern Gulf during the Pleistocene. Another possibility is that recolonization of the northern Gulf populations that we sampled may have occurred predominantly from the western Gulf, and a zone of secondary contact with more southeastern populations occurs between Mobile Bay and Tampa Bay. The Florida populations appear to be more highly differentiated from the other Gulf populations, above what is expected by distance alone, suggesting the presence of a transition zone in that region. More continuous sampling is needed to determine whether there is a zone of secondary contact in this region consistent with the southern refugia hypothesis.

The decrease in allelic richness with increasing latitude in the Gulf and the fact that allelic richness is relatively higher in the southwestern and southeastern Gulf compared to the northern populations is consistent with recolonization of the northern Gulf from southern refugia. A decrease in genetic diversity with latitude is usually hypothesized to result from the recolonization of northern habitats after the retreat of the glaciers during the Pleistocene and has been described for a number of taxa (Hewitt 2000, Adams et al. 2006, Haney et al. 2007).

Alternatively, the observed genetic discontinuities in *Fundulus grandis* may simply be located in areas that either have geographic features or selective regimes that limit dispersal and are not indicative of historical transition zones. This may be especially true of the Mobile Bay area that was relatively less differentiated from the western populations. Large bays along the east coast of North America are believed to be effective dispersal barriers for *F. heteroclitus*, which prefer shallow, near-shore water (Duvernell et al. 2008), and may explain the genetic differentiation of *F. grandis* located at Mobile Bay and Tampa Bay. It is currently unknown what potential selective factors might limit gene flow in this species, although adaptations to differing levels of hypoxia, salinity, and temperature are possibilities. A finer scale genetic study of the Mobile and Tampa Bay areas may reveal the nature of the dispersal barriers in these regions (e.g. Duvernell et al. 2008). Recent disturbances such as hurricanes could also affect the population genetic structure of *F. grandis* by altering the distribution of suitable habitat or by killing large numbers of individuals, resulting in decreased population sizes or local extinction and recolonization events. We did not find evidence of recent bottlenecks or founder events in either Mobile Bay or Florida. The effects of hurricanes on near-shore fish populations are believed to be relatively transitory (Greenwood et al. 2006).

The estimation of gene flow between populations using estimators such as F_{ST} assumes the presence of migration–drift equilibrium (e.g. Hutchison & Templeton 1999, Castric & Bernatchez 2003, Austin et al. 2004). Populations with a stepping stone population structure are expected to exhibit a pattern of IBD when they are at migration–drift equilibrium (Kimura & Weiss 1964, Crow & Aoki 1984, Slatkin 1993, Hutchison & Templeton 1999). The relatively strong IBD relationship ($r = 0.89$) and a y -intercept near the origin (0) among the populations west of Mobile Bay suggest these populations may be at migration–drift equilibrium. The genetic discontinuities revealed in Mobile Bay and especially Florida suggest that populations located along the entire Gulf coast are probably not at migration–drift equilibrium; therefore, inferences about the effects of gene flow and selection on population structure must take these patterns into consideration.

As expected, patterns of nuclear microsatellite variation suggest that contemporary gene flow is limited in *Fundulus grandis* and occurs mainly between neighboring sites. Indications of historical influences on the current day genetic structure of *F. grandis* include a negative relationship between genetic diversity and latitude and evidence of genetic discontinuities at hypothesized biogeographic boundaries, although these discontinuities may be related to current day dispersal barriers. Genotyping these populations with a mitochondrial DNA marker and extending the sampling range to include populations from more southern latitudes (i.e. extreme southern and eastern Florida and northern Mexico) could improve our understanding of the patterns we have found in the present study. Although not as dramatic as the Atlantic versus Gulf split seen in many species, results suggest that microsatellite markers can uncover subtle evidence for historical influences and genetic discontinuities within the Gulf of Mexico. More intraspecific studies are needed on other estuarine-dependent species within the Gulf to test the generality of the southern refugia hypothesis.

Acknowledgements. We thank M. Vaughn for collecting samples and J. VanWye for laboratory assistance. This research was supported by NSF BE/GEN-EN 0221879 and NIH/NHLBI HL65470. We also thank A. Hale and 4 anonymous reviewers for commenting on earlier versions of this manuscript and J. Kool for providing Fig. 1.

LITERATURE CITED

- Adams SM, Oleksiak MF, Duvernell DD (2005) Microsatellite primers for the Atlantic coastal killifish, *Fundulus heteroclitus*, with applicability to related *Fundulus* species. *Mol Ecol Notes* 5:275–277
- Adams SM, Lindmeier JB, Duvernell DD (2006) Microsatellite

- analysis of the phylogeography, Pleistocene history and secondary contact hypotheses for the killifish, *Fundulus heteroclitus*. Mol Ecol 15:1109–1123
- Austin JD, Loughheed SC, Boag PT (2004) Controlling for the effects of history and nonequilibrium conditions in gene flow estimates in northern bullfrog (*Rana catesbeiana*) populations. Genetics 168:1491–1506
- Avise JC (1996) Toward a regional conservation genetics perspective: phylogeography of faunas in the southeastern United States. In: Avise JC, Hamrick JL (eds) Conservation genetics. Chapman & Hall, New York, p 431–470
- Avise JC (2000) Phylogeography. Harvard University Press, Cambridge, MA
- Avise JC, Reeb CA, Saunders NC (1987) Geographic population structure and species differences in mitochondrial DNA of mouthbrooding marine catfishes (Ariidae) and demersal spawning toadfishes (Batrachoididae). Evolution 41:991–1002
- Ball AO, Chapman RW (2003) Population genetic analysis of white shrimp, *Litopenaeus setiferus*, using microsatellite genetic markers. Mol Ecol 12:2319–2330
- Barnwell FH, Thurman CL (1984) Taxonomy and biogeography of the fiddler crabs (Ocypodidae, genus *Uca*) of the Atlantic and Gulf coasts of eastern North America. Zool J Linn Soc 81:23–87
- Beheregaray LB, Sunnucks P (2001) Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. Mol Ecol 10:2849–2866
- Bert TM (1986) Speciation in western Atlantic stone crabs (genus *Menippe*): the role of geological processes and climatic events in the formation and distribution of species. Mar Biol 93:157–170
- Bilodeau AL, Felder DL, Neigel JE (2005) Population structure at two geographic scales in the burrowing crustacean *Callichirus islagrande* (Decapoda, Thalassinidea): historical and contemporary barriers to planktonic dispersal. Evolution 59:2125–2138
- Bilton DT, Paula J, Bishop JDD (2002) Dispersal, genetic differentiation and speciation in estuarine organisms. Estuar Coast Shelf Sci 55:937–952
- Blandon LR, Ward R, King TL, Karel WJ, Monaghan JP (2001) Preliminary genetic population structure of southern flounder, *Paralichthys lethostigma*, along the Atlantic Coast and Gulf of Mexico. Fish Bull 99:671–678
- Bloom AL (1983) Sea level and coastal morphology of the United States through the late Wisconsin glacial maximum. In: Porter SC (ed) The Late Pleistocene, Vol 1. University of Minnesota, Minneapolis, MN, p 215–229
- Briggs JC (1974) Marine zoogeography. McGraw-Hill, New York
- Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Mol Ecol 5:453–455
- Brunner CA (1982) Paleooceanography of surface waters in the Gulf of Mexico during the late Quaternary. Quat Res 17:105–119
- Burridge CP, Hurt AC, Farrington LW, Coutin PC, Austin CM (2004) Stepping stone gene flow in an estuarine-dwelling sparid from south-east Australia. J Fish Biol 64:805–819
- Castric V, Bernatchez L (2003) The rise and fall of isolation by distance in the anadromous brook charr (*Salvelinus fontinalis* Mitchell). Genetics 163:983–996
- Chenoweth SF, Hughes JM (2003) Oceanic interchange and nonequilibrium population structure in the estuarine dependent Indo-Pacific tasselfish, *Polynemus sheridani*. Heredity 80:187–197
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014
- Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. Genetics 153:1989–2000
- Crow JF, Aoki K (1984) Group selection for a polygenic behavioral trait: estimating the degree of population subdivision. Proc Natl Acad Sci USA 81:6073–6077
- Durand JD, Tine M, Panfili J, Thiaw OT, Lae R (2005) Impact of glaciations and geographic distance on the genetic structure of a tropical estuarine fish, *Ethmalosa fimbriata* (Clupeidae, S. Bowdich, 1825). Mol Phylogenet Evol 36:277–287
- Duvernell DD, Lindmeier JB, Faust KE, Whitehead A (2008) Relative influences of historical and contemporary forces shaping the distribution of genetic variation in the Atlantic killifish, *Fundulus heteroclitus*. Mol Ecol 17:1344–1360
- Engle VD, Summers JK (2000) Biogeography of benthic macroinvertebrates in estuaries along the Gulf of Mexico and western Atlantic coasts. Hydrobiologia 436:17–33
- Estoup A, Jarne P, Cornuet JM (2002) Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620
- Felder DL, Staton JL (1994) Genetic differentiation in trans-Floridian species complexes of *Sesarma* and *Uca* (Decapoda, Brachyura). J Crustac Biol 14:191–209
- Gold JR, Richardson LR (1998) Mitochondrial DNA diversification and population structure in fishes from the Gulf of Mexico and western Atlantic. J Hered 89:404–414
- Gold JR, Turner TF (2002) Population structure of red drum (*Sciaenops ocellatus*) in the northern Gulf of Mexico, as inferred from variation in nuclear-encoded microsatellites. Mar Biol 140:249–265
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available at: <http://www2.unil.ch/popgen/softwares/fstat.htm>
- Goudet J, Raymond M, deMeeus T, Rousset F (1996) Testing differentiation in diploid populations. Genetics 144:1933–1940
- Greeley MS, MacGregor R III (1983) Annual and semilunar reproductive cycles of the Gulf killifish, *Fundulus grandis*, on the Alabama Gulf coast. Copeia 1983:711–718
- Greenwood MFD, Stevens PW, Matheson RE Jr (2006) Effects of the 2004 hurricanes on the fish assemblages in two proximate southwest Florida estuaries: change in the context of interannual variability. Estuar Coasts 29:985–996
- Grosberg R, Cunningham CW (2001) Genetic structure in the sea: from populations to communities. In: Bertness MD, Gaines S, Hay ME (eds) Marine community ecology. Sinauer Associates, Sunderland, MA, p 61–84
- Gurgel CFD, Fredericq S, Norris JN (2004) Phylogeography of *Gracilaria tikvahiae* (Gracilariaceae, Rhodophyta): a study of genetic discontinuity in a continuously distributed species based on molecular evidence. J Phycol 40:748–758
- Haney RA, Silliman BR, Fry AJ, Layman CA, Rand DM (2007) The Pleistocene history of the sheepshead minnow (*Cyprinodon variegatus*): non-equilibrium evolutionary dynamics within a diversifying species complex. Mol Phylogenet Evol 43(3):743–754
- Harrison JS (2004) Evolution, biogeography, and the utility of mitochondrial 16s and COI genes in phylogenetic analysis

- of the crab genus *Austinia* (Decapoda: Pinnotheridae). *Mol Phylogenet Evol* 30:743–754
- Heist EJ, Gold JR (2000) DNA microsatellite loci and genetic structure of red snapper in the Gulf of Mexico. *Trans Am Fish Soc* 129:469–475
- Herke SW, Foltz DW (2002) Phylogeography of two squid (*Loligo pealei* and *L. plei*) in the Gulf of Mexico and north-western Atlantic Ocean. *Mar Biol* 140:103–115
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913
- Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* 53:1898–1914
- Imbrie J, McIntyre A, Moore TCJ (1983) The ocean around North America at the last glacial maximum. In: Porter SC (ed) *The Late Pleistocene*, Vol 1. University of Minnesota, Minneapolis, MN, p 230–238
- Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49:561–576
- Lee D, Gilbert CR, Hocutt CH, Jenkins RE, McAllister DE, Stauffer JR Jr (1980) *Atlas of North American freshwater fishes*. North Carolina State Museum of Natural History, Raleigh
- Lemaire C, Allegrucci G, Naciri M, Bahri-Sfar L, Kara H, Bonhomme F (2000) Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*)? *Mol Ecol* 9:457–467
- Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *J Hered* 89:238–247
- McClure MR, McEachran JD (1992) Hybridization between *Priodontes alatus* and *P. paralatus* in the Northern Gulf of Mexico (Pisces, Triglididae). *Copeia* 1992:1039–1046
- McIntyre A, Kipp AG, Be AWH, Crowley T and others (1976) Glacial North America 18,000 years ago: a climate reconstruction. In: Cline RM, Hays JD (ed) *Late Quaternary paleoceanography and paleoclimatology*. Geological Society of America, Boulder, CO, p 43–76
- Pampoulie C, Gysels ES, Maes GE, Hellemans B, Leentjes V, Jones AG, Volckaert FAM (2004) Evidence for fine-scale genetic structure and estuarine colonisation in a potential high gene flow marine goby (*Pomatoschistus minutus*). *Heredity* 92:434–445
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* 90:502–503
- Piry S, Alapetite A, Cornuet JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: a software for genetic assignment and first-generation migrant detection. *J Hered* 95:536–539
- Pritchard JK, Wen W (2003) Documentation for structure software: version 2.2. 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
- Purcell JFH, Cowen RK, Hughes CR, Williams DA (2006) Weak genetic structure indicates strong dispersal limits: a tale of two coral reef fish. *Proc R Soc Lond B* 273:1483–1490
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci USA* 94:9197–9201
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and eucumenicism. *J Hered* 86:248–249
- Reeb CA, Avise JC (1990) A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, *Crassostrea virginica*. *Genetics* 124:397–406
- Skinner MA, Courtenay SC, Parker WR, Curry RA (2005) Site fidelity of mummichogs (*Fundulus heteroclitus*) in an Atlantic Canadian estuary. *Water Qual Res J Can* 40:288–298
- Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47:264–279
- Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. *Mol Ecol* 15:4261–4293
- Stanley SM (1986) Anatomy of a regional mass extinction. Plio-Pleistocene decimation of the western Atlantic bivalve fauna. *Palaios* 1:17–36
- Sweeney J, Deegan L, Garritt R (1998) Population size and site fidelity of *Fundulus heteroclitus* in a macrotidal salt-marsh creek. *Biol Bull* 195:238–239
- 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
- Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J Hered* 89:438–450
- Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, fresh-water, and anadromous fishes. *J Fish Biol* 44:213–232
- Watts RJ, Johnson MS (2004) Estuaries, lagoons and enclosed embayments: habitats that enhance population subdivision of inshore fishes. *Mar Freshw Res* 55:641–651
- Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* 38:1358–1370
- Whitehead A, Crawford DL (2006) Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci USA* 103:5425–5430

Editorial responsibility: Jana Davis,
Annapolis, Maryland, USA

Submitted: March 7, 2008; Accepted: September 15, 2008
Proofs received from author(s): December 9, 2008