Genetic Population Structure of US Atlantic Coastal Striped Bass (Morone saxatilis)

David T. Gauthier
Old Dominion University, dgauthie@odu.edu

Corinne A. Audemard

Jeanette E. L. Carlsson

Tanya L. Darden

Michael R. Denson

See next page for additional authors

Follow this and additional works at: https://digitalcommons.odu.edu/biology_fac_pubs

Part of the Ecology and Evolutionary Biology Commons, Genetics Commons, and the Marine Biology Commons

Original Publication Citation

This Article is brought to you for free and open access by the Biological Sciences at ODU Digital Commons. It has been accepted for inclusion in Biological Sciences Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
Authors
David T. Gauthier, Corinne A. Audemard, Jeanette E. L. Carlsson, Tanya L. Darden, Michael R. Denson, Kimberly S. Reece, and Jens Carlsson
Genetic Population Structure of US Atlantic Coastal Striped Bass (Morone saxatilis)

DAVID T. GAUTHIER, CORINNE A. AUDEMARD, JEANETTE E. L. CARLSSON, TANYA L. DARDEN, MICHAEL R. DENSON, KIMBERLY S. REECE, AND JENS CARLSSON

From the Department of Biological Sciences, Old Dominion University, 202E Mills Godwin Building, Norfolk, VA 23529 (Gauthier); Virginia Institute of Marine Science, The College of William and Mary, PO Box 1346, Gloucester Point, VA (Audemard and Reece); Duke University, Duke Marine Lab, 135 Duke Marine Lab Road, Beaufort, NC (Jeanette Carlsson and Jens Carlsson); Marine Resources Research Institute, South Carolina Department of Natural Resources, Hollings Marine Laboratory, 331 Fort Johnson Road, Charleston, SC (Darden and Denson); School of Biological, Earth and Environmental Sciences/Aquaculture and Fisheries Development Centre, University College Cork, Distillery Fields, North Mall, Cork, Ireland (Jens Carlsson); and School of Biology and Environmental Science, UCD Science Education and Research Centre – West, University College Dublin, Dublin 4, Ireland (Jens Carlsson).

Address correspondence to D. T. Gauthier at the address above, or e-mail: dgauthie@odu.edu.

Data deposited at Dryad: http://dx.doi.org/10.5061/dryad.vp4n8

Abstract

Genetic population structure of anadromous striped bass along the US Atlantic coast was analyzed using 14 neutral nuclear DNA microsatellites. Young-of-the-year and adult striped bass (n = 1114) were sampled from Hudson River, Delaware River, Chesapeake Bay, North Carolina, and South Carolina. Analyses indicated clear population structure with significant genetic differentiation between all regions. Global multilocus $F_{ST}$ was estimated at 0.028 ($P < 0.001$). Population structure followed an isolation-by-distance model and temporal sampling indicated a stable population structure more than 2 years at all locations. Significant structure was absent within Hudson River, whereas weak but significant genetic differences were observed between northern and southern samples in Chesapeake Bay. The largest and smallest effective striped bass population sizes were found in Chesapeake Bay and South Carolina, respectively. Coalescence analysis indicated that the highest historical gene flow has been between Chesapeake Bay and Hudson River populations, and that exchange has not been unidirectional. Bayesian analysis of contemporary migration indicated that Chesapeake Bay serves as a major source of migrants for Atlantic coastal regions from Albemarle Sound northward. In addition to examining population genetic structure, the data acquired during this project were capable of serving as a baseline for assigning fish with unknown origin to source region.

Key words: anadromy, finfish, microsatellites, Moronid, population genetics

Introduction

The anadromous striped bass (Morone saxatilis; Walbaum 1792) is a dominant piscivorous fish in estuaries along the US Atlantic coast (Hartman and Brandt 1995) and fills a critical ecological niche in estuarine food webs. Atlantic coastal striped bass are present from the St Lawrence River in Canada to the St Johns River in Florida (Murdy et al. 1997). This species is also found in the Gulf of Mexico and has been introduced to inland lakes and reservoirs, as well as the US Pacific coast. Most Atlantic adult striped bass north of Albemarle Sound, North Carolina (NC), are migratory (Boreman and Lewis 1987), whereas more southern rivers are considered to harbor largely nonmigratory populations (Greene et al. 2009). Multiple life-history patterns are employed by striped bass in major riverine and estuarine systems harboring migratory populations, with portions of the population displaying varying degrees of residence and migration, including year-round residence in freshwater (Morris et al. 2003; Zlokovitz et al. 2003; Secor and Piccoli 2007). Spawning occurs in freshwater portions of tributaries throughout this species’ range; however, Chesapeake Bay (CB) is traditionally considered to be...
the major coastal production area (Berggren and Lieberman 1978; Van Winkle et al. 1988).

Striped bass is one of the most economically important finfish species along the Atlantic coast and has historically experienced considerable fishing pressure. Commercial and recreational landings declined precipitously during the 1970s and 1980s, leading to development of stringent fishing regulations by the Atlantic States Marine Fisheries Commission (1981; Weaver et al. 1986). Strict harvest laws and high recruitment during the late 1980s and early 1990s coincided with a rebound of striped bass numbers. A limited fishery reopened during 1990, with fishing restrictions further relaxed in 1995 as the census numbers had recovered (Field 1997; Richards and Rago 1999). Because of the popularity of striped bass among commercial and recreational fishers, continued effective management will be crucial in perpetuating the success of the resource. Information about how striped bass populations are structured in space and time, as well as the level of connectivity among populations is essential for such management efforts. Detection of biologically isolated populations could allow for regulation of individual management units (MUs; Moritz 1994), thus optimizing conservation of individual Atlantic coast populations.

Striped bass population structure has been addressed by previous studies at various spatial scales and with a range of genetic markers. Although genetic differentiation of geographically widely separated populations has been demonstrated (Wirgin et al. 1989, 1993; Diaz et al. 1997), no work has yet comprehensively addressed major production areas along the Atlantic coast. Further, genetic structure of striped bass within important spawning estuaries, most notably between estuaries within CB, has been a source of perennial debate (see, e.g., Brown et al. 2005). Current age-structured stock assessments of striped bass (Atlantic States Marine Fisheries Commission 2011) do not take into account genetically differentiated populations of these fish in this region. Although major production areas are assessed separately (e.g., Hudson River [HR] and CB), our knowledge of the contribution of these areas to the migratory stock is limited to studies performed before major stock collapses in the 1980s and 1990s and without the benefit of modern molecular tools (Berggren and Lieberman 1978; Van Winkle et al. 1988). Measurements of population-specific recruitment to the migratory stock may be performed via assignment testing or mixed-stock analysis of migratory adults; however, these tests are dependent on adequate genetic baseline data from the major production areas. In this work, we use 14 microsatellite loci to assess genetic population structure and demographies of the striped bass both within CB and along the Atlantic coast, including HR, Delaware River (DR), Albemarle Sound (Roanoke River), and South Carolina (SC). We assess the temporal stability of population structure in these production areas over the course of 2 consecutive years (2008–2009), and perform a preliminary test of the capability of these baseline data to allow assignment of adult fish from mixed samples to natal population.

### Materials and Methods

#### Samples

Young-of-the-year (YOY) striped bass were collected from the following locations during the summer and fall of 2008 and 2009: HR (2 sites), DR (1 site), CB (5 sites), NC (2 adjacent sites), and SC (2 sites). YOY striped bass were also collected from the York and Rappahannock Rivers within CB in 2006 and 2007, respectively (Table 1 and Figure 1). As discussed by others (see, e.g., Brown et al. 2005), use of YOY and temporal replicates avoids potential artifacts inherent in data sets using adult fishes, most notably those due to dispersal, differential structure of age classes, and sex ratio biases. In 2008, scale samples from spawning adults in HR were added to baseline data from this location due to degradation of some YOY samples during shipment. SC tissue samples from 2007 were obtained from adults in spawning condition from the Santee–Cooper River system. Although these HR 2008 and SC 2007 samples do not represent a direct measurement of YOY genetic structure, such structure should be reflected in

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Striped bass sample regions, locations, year of sampling and number of fish sampled (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
<td>Location</td>
</tr>
<tr>
<td>Hudson River (HR)</td>
<td>Upper* (RK140-147)</td>
</tr>
<tr>
<td></td>
<td>Upper* (RK140-185)</td>
</tr>
<tr>
<td></td>
<td>Upper (RR90)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower (RK56)</td>
</tr>
<tr>
<td>Delaware river (DR)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
</tr>
<tr>
<td>Chesapeake Bay (CB)</td>
<td>Head of Bay</td>
</tr>
<tr>
<td></td>
<td>Potomac</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rappahannock</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>York (MP)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>York (PK)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>James</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>North Carolina (NC)</td>
<td>BW</td>
</tr>
<tr>
<td></td>
<td>EB</td>
</tr>
<tr>
<td></td>
<td>BW</td>
</tr>
<tr>
<td>South Carolina (SC)</td>
<td>Santee–Cooper</td>
</tr>
<tr>
<td></td>
<td>Lake Moultrie</td>
</tr>
<tr>
<td></td>
<td>Lake Moultrie</td>
</tr>
<tr>
<td></td>
<td>Lake Marion</td>
</tr>
</tbody>
</table>

MP, Mattaponi tributary of York River; PK, Pamunkey tributary of York River; BW, Black Walnut Point; EB, Edenhouse Bridge. Hudson River locations include river kilometer (RK) designations.

* Scale samples from adult male fish.

† Scale samples from adult female fish.

‡ Tissue samples (fin clip) from male and female adult fish.
Figure 1. Sampling locations (2008–2009) for YOY striped bass are shown in circles: 2008 (open circles), 2009 (closed circles), 2008–2009 (half-closed circles). Locations of scale collections (Hudson River 2008 only) are shown as closed squares. Boxes in overview map (upper left) show major regions detected by STRUCTURE analysis, including Hudson River (a), Delaware River + Chesapeake Bay (b), North Carolina (c), and South Carolina (d).

data from male and female spawning adults. For elucidating the potential of assignment testing, age 1+ fish (n = 55) were collected by the Chesapeake Bay Multispecies Monitoring and Assessment Program (ChesMMAP) (Latour et al. 2008), and larger adults (n = 76) were collected from near-offshore waters along the East coast from Cape May, NJ to Long Island Sound by the Northeast Area Monitoring Assessment Program (NEAMAP) survey.
Molecular Markers

DNA was extracted from fin clips or scales with either the DNassets Tissue Kit (Qiagen Inc., Santa Clara, CA) or by proteinase K/chelex extraction (Estoup et al. 1996). Two groups of YOY striped bass were used for initial microsatellite screening, comprising 21 individuals collected from the York and James River tributaries of CB in 2006 and 48 individuals obtained in 2007 from the Santee–Cooper river system in SC. Approximately 30 markers were screened in this work from published striped bass microsatellite loci (Couch et al. 2006; Rexroad et al. 2006). The microsatellite multiplex panels used in this study were modifications of panels from previously published work (Fountain et al. 2009) and comprised 14 loci: 9 from Couch et al. (2006) and 5 unpublished markers deposited in GenBank (see Supplementary Table S1 online). These loci were PCR-amplified and screened for allelic variability using either ABI 3130xl or 3730xl Genetic Analyzers (Applied Biosystems, Forest City, CA), with identical chemistries for both instruments. The size of individual microsatellite alleles, in base pairs, was measured with GENEMARKER software (SoftGenetics, State College, PA), using GeneScan™ 500 LIZ® size standard. Approximately 20% of all the samples were rerun to assess repeatability of allele scoring by a given observer, and a total of 225 genotypes were rerun to assess repeatability in scoring between sequencing platforms. Consistency of scoring exceeded 99% in the former case and was 100% in the latter.

Statistical Approaches

Quality control analyses for microsatellites were performed with MICRO-CHECKER v.2.2.3 software (van Oosterhout et al. 2004). GENEPOP v. 4.0.9 software (Rousset 2008) was used to analyze allele frequencies for deviations from Hardy–Weinberg expectations (HWE, exact test; Guo and Thompson 1992), create estimations of observed (\(H_{\text{obs}}\)) and expected (\(H_{\text{exp}}\)) heterozygosities, test for heterozygosity excess or deficiency, and test for linkage disequilibrium among loci (exact tests). Samples from SC showed gametic phase disequilibrium (see Results); therefore, the presence of full-sibs was investigated by applying the maximum likelihood method implemented in COLONY v. 2.0.2.1 (Jones and Wang 2010). The analysis was based on the full-likelihood method and the “short length of run,” “full-likelihood,” and “medium likelihood precision.” Marker error rates were set to 0.1% for all loci. Allelic richness was estimated with the FSTAT v. 2.9.3 software (Goudet 1995). LOSITAN software (Beaumont and Nichols 1996; Antao et al. 2008) was used to detect evidence for selection at individual loci. The analysis of molecular variance (AMOVA) algorithm included in the computer program package ARLEQUIN v. 3.5 (Excoffier and Lischer 2010) was used to partition the total observed variance into between-years variability (to estimate the temporal stability), and between-locations variability (to estimate the spatial structure) as well as compare the variability among regions (regional variability).

The software MICRO-SATELLITE ANALYZER (MSA; Dieringer and Schlötterer 2003) was used to estimate and test significance of Weir and Cockerham’s (1984) unbiased estimator of Wright’s \(F\)-statistics, \(F_{ST}\) (10,000 permutations).

BOTTLENECK software (Cornuet and Luikart 1996) was used to determine if data suggested the occurrence of severe past reductions in striped bass effective population size. The program MIGRATE-N v. 3.1.6 (Beerli and Felsenstein 2001) was used to estimate theta (\(\Theta\)), which is equal to \(4N_e\mu\) where \(N_e\) is the long-term (inbreeding) effective population size and \(\mu\) is the mutation rate. When inspecting how the effective population size is estimated (\(N_e = \Theta/4\mu\)), it is evident that the mutation rate has a large influence on the estimate. The reported mutation rates (\(\mu\)) for microsatellites range from \(10^{-5}\) to \(10^{-2}\) per locus per generation (Weber and Wong 1993) and consequently, any estimate of effective population size is prone to very large variation depending on which mutation rate is used. The mutation rate should, however, be the same for each marker independent of the population studied. For this reason, no attempt was made to estimate the actual effective population size; instead the \(\Theta\) estimates were used as proxies for relative effective population sizes. Simultaneously, MIGRATE-N was used to estimate \(M\), where \(M = m/\mu\), \(m\) is the migration rate per generation and \(\mu\) is the mutation rate. The number of immigrants per generation, \(4N_e\mu\) (for nuclear data), can be estimated by multiplying \(\Theta\) by \(M\). The software BAYESASS, v. 3.0.1 (Wilson and Rannala 2003) was used to estimate contemporary gene flow among populations. Both MIGRATE-N and BAYESASS runs were performed under a 5-population (HR, DR, CB, NC, and SC) scenario as indicated by \(F_{ST}\) analysis (see Results), and results of 5 runs were averaged. Convergence of models in BAYESASS was examined by comparison of 5 runs with different random starting seeds, as well as plotting total log-likelihood score versus iteration with TRACER v. 1.5 (Rambaut and Drummond 2007). Twenty-one million iterations with a 2 million iteration burn-in were used in BAYESASS. STRUCTURE v. 2.3.3 software (Pritchard et al. 2000; Falush et al. 2007) was used to sort individuals into clusters (using the admixture and correlated allele frequency model with 1000000 replicates and a burn-in length of 100000) using \(K\) (number of clusters) from one to five with 10 replicates for each \(K\). Following recommendations by Hubisz et al. (2009), when population structure is weak, STRUCTURE was implemented with and without the “use sampling locations as prior (LOCPRIOR)” switch. The most likely \(K\) was assessed by plotting Ln(PD) and implementing the \(\Delta K\) method as in Evanno et al. (2005). The accuracy of STRUCTURE-based assignments of individuals to the correct source population was estimated by using the option “use populating information to test for migrants” for the potential source populations; this option was turned off for fish that were being assigned. Ten individuals per cluster (HR, DR/CB, NC, and SC, see below) were removed from potential source populations and assigned back to source population. The software CLUMPP v. 1.1.2 (Jakobsson
was used to account for cluster label switching and assign clusters to which each run corresponded (search options: greedy, G′, using random input orders and 1000 repeats). Self-assignment tests (i.e., testing whether individual YOY striped bass from a specific population were correctly assigned to their population of origin) were performed using GENECLASS2 software (Piry et al. 2004) with the Bayesian method of Rannala and Mountain (1997). The sequential Bonferroni technique (Rice 1989) was used to adjust significance levels in cases with multiple tests.

Results

Genetic Variability

Quality control with MICRO-CHECKER v. 2.2.3 indicated that allele scoring was not affected by technical artifacts, stutter, or large allele drop-outs (1000 randomizations). Locus MSM1603 showed potential effects of null alleles in HR and CB samples, and MSM1094 showed indication of null alleles in SC. There were no consistent, across-samples effects of null alleles, and given that null alleles only have minor effects on \( F_{ST} \) estimates and the accuracy of assignment testing (Carlsson 2008), these loci were included in the downstream analyses. All samples were screened at a total of 14 microsatellite loci. Only individuals for which at least 10 loci could be scored were included in statistical analyses. Summary statistics for the 5 populations indicated by \( F_{ST} \) analysis are presented in Supplementary Table S2 online. Indications of linkage disequilibrium were tested within each of the 5 regions (HR, DR, CB, NC, and SC); a total of 64 locus pairs initially indicated linkage but only 13 combinations remained significant after correction for multiple tests (sequential Bonferroni correction, \( k = 455 \)). Linkage was not indicated for any locus combination in all 5 regions, and most linked combinations were found among the SC samples. Therefore, it is likely that the loci were not actually physically linked, and that gametic phase disequilibrium may account for the data from the SC sample. Further analysis of the SC samples with COLONY indicated that the 106 individuals from SC comprised 72 full-sib families, with family size ranging from 1 to 8 full-sibs, with 11 families consisting of at least 3 individuals, supporting this hypothesis. Similar results have been found in Santee–Cooper striped bass cohorts by Liu and Ely (2009). No loci deviated from HWE across all sample locations. No evidence for selection was found at any locus via LOSITAN analysis under either infinite allele or stepwise mutation models.

Temporal Structure

Prior to further analyses, temporal variability between annual replicates within a geographic location was assessed. There was no significant multilocus \( F_{ST} \) structure (\( P > 0.05 \) for all tests) between temporal replicates at any site; consequently, these replicates were pooled within-site for further analysis. AMOVA analysis using only samples with \( n > 29 \) indicated that temporal variability (\( F_{ST} \)) was negligible (\( F_{SC} = -0.002; \ P > 0.05 \)) compared with spatial variability (\( F_{CT} \)), which was considerable (\( F_{CT} = 0.021; \ P < 0.001 \)).

Spatial Structure

\( F_{ST} \) analyses indicated significant population structure between the 5 study regions. Global multi- and single-locus \( F_{ST} \) were estimated at 0.028 (\( P < 0.001 \)) and 0.014–0.062, respectively (Supplementary Table S2 online). Pairwise multilocus \( F_{ST} \) estimates were all significantly different from zero after Bonferroni correction and ranged from 0.002 between the DR and CB samples to 0.089 between the HR and SC samples (Table 2). The degree of genetic differentiation (\( F_{ST} \)) increased with increasing distance between river mouths for all regions aside from SC, which was equally divergent from all other regions (Mantel test, \( P = 0.012 \)) indicating an isolation-by-distance model. Because of the possibility that SC would disproportionately affect this analysis due to the presumed nonmigratory nature of its striped bass population, it was removed and a second analysis was performed on the remaining regions. Isolation-by-distance model continued to be supported for the remaining 4 populations (\( P = 0.036 \)). Analysis of STRUCTURE output by plotting Ln(PD) showed equal support for sorting of genotypes into 2 or 3 clusters (Supplementary Figure S3 online). When using the LOCPRIOR, however, Ln(PD) indicated 4 clusters (i.e., HR, DR/CB, NC, and SC). As the SC sample clearly showed the largest genetic differentiation, STRUCTURE analysis was performed without the SC sample using LOCPRIOR. The resulting Ln(PD) plot indicated 3 clusters.

Within HR, significant differences in multilocus \( F_{ST} \) estimates were detected neither between the scale samples from

Table 2 Multilocus pairwise estimates among regional samples

<table>
<thead>
<tr>
<th>Regional</th>
<th>HR</th>
<th>DR</th>
<th>CB</th>
<th>NC</th>
<th>SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>DR</td>
<td>0.0101</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CB</td>
<td>0.0086</td>
<td>0.0022</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>0.0122</td>
<td>0.0089</td>
<td>0.0059</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td>0.0885</td>
<td>0.0867</td>
<td>0.0853</td>
<td>0.087</td>
<td></td>
</tr>
</tbody>
</table>

Chesapeake Bay

<table>
<thead>
<tr>
<th></th>
<th>HB</th>
<th>PT</th>
<th>RP</th>
<th>YK</th>
<th>JM</th>
</tr>
</thead>
<tbody>
<tr>
<td>HB</td>
<td>0.5954</td>
<td>0.1779</td>
<td>0.351</td>
<td>0.0091</td>
<td></td>
</tr>
<tr>
<td>PT</td>
<td>0.0003</td>
<td>0.0383</td>
<td>0.0866</td>
<td>0.339</td>
<td></td>
</tr>
<tr>
<td>RP</td>
<td>0.0006</td>
<td>0.0015</td>
<td>0.802</td>
<td>0.0564</td>
<td></td>
</tr>
<tr>
<td>YK</td>
<td>0.0003</td>
<td>0.0013</td>
<td>0.0007</td>
<td>0.1599</td>
<td></td>
</tr>
<tr>
<td>JM</td>
<td>0.0024</td>
<td>0.0003</td>
<td>0.001</td>
<td>0.0009</td>
<td></td>
</tr>
</tbody>
</table>

HB, Head of Bay; JM, James; PT, Potomac; RP, Rappahannock; YK, York.
females and males nor between the lower and upper HR ($P > 0.05$); therefore, all HR samples were pooled within year. No structure was evident between the 2 NC locations ($P > 0.05$), as expected due to their close proximity (<7 km), therefore these were similarly pooled. The global multilocus $F_{ST}$ for CB was estimated at 0.001 ($P = 0.036$) and indicated shallow, but significant population structure among striped bass in this region. The structure did not follow an isolation-by-distance pattern (Mantel test, $P > 0.05$). Pairwise comparisons between striped bass from the 5 sampling localities in CB were not significantly different from zero after Bonferroni correction ($P > 0.005$ for all tests, $k = 10$). Pooling of the 2 most northern (Head of Bay and Potomac River) and 3 most southern regions (Rappahannock, York, and James Rivers) gave a significant multilocus $F_{ST}$ estimate (0.001, $P = 0.002$) between northern and southern groups, indicating substructure. However, the level of intra-CB differentiation was much weaker (~30x) than regional structure (see above) and CB samples were therefore pooled for further analyses.

### Effective Population Size

Long-term relative effective population size ($\Theta$ as estimated by MIGRATE-N) indicated the highest values in CB ($\Theta = 2.35$) and the lowest values in SC ($\Theta = 1.42$; Table 3). None of the 5 identified populations showed significant reductions in effective population size via the BOTTLENECK analysis (1-tailed Wilcoxon test of heterozygosity excess, all $P > 0.05$) using the 2-phase mutation-drift equilibrium model, variance = 10% and $P = 90\%$).

### Contemporary and Historical Gene Flow

Contemporary gene flow was analyzed with BAYESASS under a 5-population scenario. No significant migration to other systems was measured from HR, DR, NC, or SC populations, whereas CB contributed significantly to all systems aside from SC. CB and SC were essentially self-recruiting (Table 3). It should be noted that this program limits total proportion of migrants in a population to one-third, therefore values of approximately 0.67 and 0.33 for nonmigration and migration, respectively, represent algorithm bounds and not necessarily actual migrant proportions. In the case of HR, migrant proportion was estimated at 0.265 (95% credible set 0.242–0.289), and is therefore not representative of a limitation in model bounds. Model diagnostics for these analyses were performed according to previous recommendations (Wilson and Rannala 2003; Faubet et al. 2007), and good model convergence was demonstrated. The BAYESASS method performs less well when $F_{ST}$ values between populations are less than 0.05, as well as when migration rates are high (Faubet et al. 2007), however, and caution is warranted in interpreting these results as exact measurements of migration. MIGRATE-N analysis of long-term gene flow demonstrated unequal migration rates among the 5 examined populations. Compared with contemporary gene flow

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Receiver</th>
<th>HR</th>
<th>DR</th>
<th>CB</th>
<th>NC</th>
<th>SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAYESASS ($N_m$)</td>
<td></td>
<td>0.727 (0.705 to 0.750)</td>
<td>0.003 (-0.002 to 0.004)</td>
<td>0.265 (0.242 to 0.289)</td>
<td>0.003 (-0.003 to 0.009)</td>
<td>0.002 (-0.002 to 0.006)</td>
</tr>
<tr>
<td></td>
<td>HR</td>
<td>0.001</td>
<td>0.669 (0.663 to 0.675)</td>
<td>0.320 (0.308 to 0.332)</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>DR</td>
<td>0.000</td>
<td>0.000</td>
<td>0.995 (0.991 to 0.999)</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>0.003</td>
<td>0.002</td>
<td>0.318</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>SC</td>
<td>0.005</td>
<td>0.003</td>
<td>0.303 (0.303 to 0.333)</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>MIGRATE-N ($\Theta_M$)</td>
<td></td>
<td>1.97 (0.12)</td>
<td>4.91 (3.73 to 6.09)</td>
<td>20.58 (17.55 to 23.39)</td>
<td>5.38 (4.20 to 6.78)</td>
<td>4.22 (3.00 to 5.90)</td>
</tr>
<tr>
<td></td>
<td>HR</td>
<td>4.93 (2.76 to 7.17)</td>
<td>2.18 (0.12)</td>
<td>11.85 (9.64 to 15.27)</td>
<td>3.68 (2.21 to 6.15)</td>
<td>3.24 (1.71 to 4.47)</td>
</tr>
<tr>
<td></td>
<td>DR</td>
<td>17.82 (15.20 to 20.37)</td>
<td>11.77 (8.97 to 14.58)</td>
<td>2.35 (0.05)</td>
<td>14.13 (12.53 to 15.89)</td>
<td>9.85 (8.34 to 11.50)</td>
</tr>
<tr>
<td></td>
<td>CB</td>
<td>4.91 (3.33 to 6.40)</td>
<td>3.74</td>
<td>15.43 (13.25 to 19.13)</td>
<td>2.04 (0.15)</td>
<td>3.05 (1.96 to 4.11)</td>
</tr>
<tr>
<td></td>
<td>NC</td>
<td>2.19 (1.65 to 2.76)</td>
<td>2.03 (1.42 to 2.82)</td>
<td>6.52 (4.53–7.55)</td>
<td>2.20 (1.48 to 2.85)</td>
<td>1.42 (0.12)</td>
</tr>
</tbody>
</table>

For BAYESASS analyses, values with a credible set (CS) excluding zero are in bold. Estimated values for both analyses are means of 5 model runs. For BAYESASS, 95% CSs are constructed as ±1.96 SDs, and the highest and lowest observed bounds of the CS among model runs are presented. For MIGRATE-N, the 95% confidence interval is constructed from minimum and maximum estimates of 0.025 and 0.975 percentiles, respectively, among runs. Long-term relative effective population size ($\Theta$) for each region as estimated by MIGRATE-N (averaged over 5 runs) is given on the identity diagonal in italics.
Table 4  Assignment testing for fish collected by ChesMMAP (within Chesapeake Bay) and NEAMAP (nearshore Atlantic, Cape May, NJ to Long Island Sound)

<table>
<thead>
<tr>
<th></th>
<th>Percentage assigned</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Scenario</td>
<td>HR</td>
<td>DR</td>
<td>CB</td>
</tr>
<tr>
<td>ChesMMAP</td>
<td>55</td>
<td>4-pop</td>
<td>100</td>
<td>12.7</td>
<td>76.3</td>
</tr>
<tr>
<td></td>
<td>5-pop</td>
<td>100</td>
<td>10.9</td>
<td>34.5</td>
<td>43.6</td>
</tr>
<tr>
<td>NEAMAP</td>
<td>76</td>
<td>4-pop</td>
<td>100</td>
<td>15.8</td>
<td>78.9</td>
</tr>
<tr>
<td></td>
<td>5-pop</td>
<td>100</td>
<td>13.2</td>
<td>31.6</td>
<td>50.0</td>
</tr>
</tbody>
</table>

80% assignment score cutoff

<table>
<thead>
<tr>
<th></th>
<th>Percentage assigned</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Scenario</td>
<td>HR</td>
<td>DR</td>
<td>CB</td>
</tr>
<tr>
<td>ChesMMAP</td>
<td>55</td>
<td>4-pop</td>
<td>56.4</td>
<td>6.4</td>
<td>90.3</td>
</tr>
<tr>
<td></td>
<td>5-pop</td>
<td>26.0</td>
<td>6.7</td>
<td>46.7</td>
<td>40.0</td>
</tr>
<tr>
<td>NEAMAP</td>
<td>76</td>
<td>4-pop</td>
<td>72.4</td>
<td>14.5</td>
<td>83.6</td>
</tr>
<tr>
<td></td>
<td>5-pop</td>
<td>25.0</td>
<td>42.1</td>
<td>26.3</td>
<td>31.6</td>
</tr>
</tbody>
</table>

Testing was performed under 4-population (as determined by STRUCTURE) and 5-population (as determined by $F_{ST}$ analysis) scenarios. DR and CB are collapsed into a single region under the 4-population scenario. The lower half of the table shows percentage of fish assigned under 80% assignment cutoff, including both total number assigned and percentage by region.

as estimated by BAYESASS, historical gene flow was much less unidirectional, with major production centers CB and HR demonstrating relatively large effective migration rates whose confidence intervals overlapped. Lesser degrees of bidirectional gene flow were observed between other regions.

Assignment Testing

GENECLASS2 was used both on the set of 5 populations identified by $F_{ST}$ analysis and on the 4 populations identified by STRUCTURE. Under the 5- and 4-population scenario, 61.9% and 77.1% of baseline samples were assigned back to their original (known) population, respectively. In the absence of population structure the expected proportion would be 20% and 25%, respectively; therefore, the GENECLASS2 results strongly corroborated those from the $F_{ST}$ analysis. Using a stringent cutoff of 80% minimum assignment score in GENECLASS2, 40.2% (5-population) and 59.9% (4-population) of baseline fish could be assigned to any population, and of these subgroups, 82.0% and 88% of fish were correctly assigned to known origin, respectively.

GENECLASS2 assignments of ChesMMAP and NEAMAP samples were similar under both 5- and 4-population scenarios, with >75% of unknowns being assigned to CB and/or DR and none assigning to SC (Table 4). Imposing an 80% assignment score cutoff sharply reduced the number of fish assigned under the 5-population scenario, whereas in the 4-population scenario, 56% and 72% of fish from ChesMMAP and NEAMAP surveys were assignable under this restriction, respectively. Note that those fish not being assignable in the previous analyses due to low assignment score (<80) may represent true F1 or F2 hybrids, or harbor genetic variation that is common in several potential source populations. Fish assigning to CB and/or DB comprised a larger percentage of total assignees under the 80% cutoff compared with no cutoff, with the exception of the NEAMAP sample under the 5-population scenario, in which

a large (42%) percentage of the assignments with strong support indicated HR origin.

Discussion

Population Structure from Hudson River to South Carolina

Although several previous studies have addressed striped bass population genetic structure on various scales and with different tools, our study is the first to use microsatellites on samples from all major production areas (HR to SC). Our analysis of genetic population structure of striped bass along the US Eastern seaboard indicated a clear genetic structure with significant population differentiation among all regions sampled (HR, DR, CB, NC, and SC). The structure followed an isolation-by-distance model with increasing genetic differentiation with increasing waterway distance among populations. All locations were sampled at least 2 years and we were unable to detect any temporal genetic variability among years, indicating that the structure is temporally stable over at least the sampling period.

Population Structure within Chesapeake Bay

Genetic population structure of striped bass within CB has been examined previously by several authors, with equivocal results. Early protein-based studies (e.g., Morgan and Koo 1973; Grove et al. 1976; Sidell et al. 1980) were limited by extremely low variability of electrophoretic profiles in this species (Waldman et al. 1988), and the issue was subsequently explored by several authors using restriction fragment length polymorphism (RFLP) of mtDNA. Some studies indicated intra-Bay structuring via this method; however, this structure did not demonstrate temporal stability (Chapman 1987) or was potentially complicated by the use of young adult fish, which may have dispersed from natal estuaries (Chapman 1990; Wirgin et al. 1990). Neither Laughlin and Turner (1996) nor Brown et al. (2005) found conclusive evidence for structure within the Bay, using variable nucleotide tandem repeat and microsatellite markers, respectively. Brown et al. (2005) also re-analyzed previously published mtDNA RFLP data for fixation indices and found that the majority of these data (Wirgin et al. 1990; Wirgin et al. 1997) did not support existence of structure in the Bay. Data of Chapman (1990) did produce $F_{ST}$ estimates significantly different from 0 for 2 annual subsets; however, as discussed by both Brown et al. (2005) and Chapman et al. (1990) this may be due to “asymmetric homing” or bias of data by differential homing of male and female fish. In this work, YOY striped bass are used to eliminate complications from migration of young adults, as juvenile striped bass tend to move downstream during their first year, but remain in or near their natal estuary until at least their second year of life (see Fay et al. 1983; Greene et al. 2009). Further, the use of microsatellite markers eliminates issues with potential asymmetric homing. In contrast with the microsatellite-based study of YOY striped bass by Brown et al. (2005), we did find evidence for
significant global population structure within CB; however, this structuring was weak ($F_{ST} = 0.001$) and pairwise comparisons between individual rivers were not significant after correction for multiple tests. The body of evidence in this area, therefore, indicates that genetic population structure of striped bass within CB is shallow and that straying of adults between spawning areas may occur.

**Contemporary and Historical Migration**

Bayesian analysis of contemporary gene flow among populations in this study reflected essentially total self-recruitment for all studied systems apart from CB. Chesapeake Bay appeared to be the only population contributing emigrants to surrounding populations, whereas fish from those populations appeared to return exclusively to their natal rivers. There were no indications that CB received reproducing immigrants from other populations. Striped bass are thought to overwinter in mixed-stock assemblages from NJ to Cape Hatteras (Waldman et al. 2012), and display a general pattern of northward movement in summer and southward movement in winter (Welsh et al. 2007). The large majority of migrating fish tagged in HR are recaptured north of Cape May, NJ although recaptures of fish tagged in HR do occur within CB (Dorazio et al. 1994) and in coastal waters off NC (Waldman et al. 1990), and fish of HR origin are found in samples within CB (this study). There is considerable evidence that fish of CB origin migrate in significant numbers to the DR and HR, and several authors have attributed abundant year classes of striped bass in northern waters to strong production in CB (see, Kohlenstein 1981). It appears from the BAYESASS-derived contemporary migration data in this study that there is not only northern movement of CB fish but also potential limited reproduction of these fish in HR, whereas the reciprocal phenomenon does not appear to occur.

Although BAYESASS analysis indicates a unidirectional gene flow from CB to HR, long-term historical analysis based on MIGRATE-N indicates a much more multidirectional gene flow over long periods of time. Therefore, over long time scales, the CB, although being the major source of migrants to other systems, has also received gene flow. Reliable historical data for striped bass migrational patterns are limited by lack of tagging studies prior to the 1930s and records of spawning prior to this time, especially in northern rivers, are largely anecdotal. From the descriptions of Merriman (1941), however, it seems likely that spawning of striped bass has historically occurred in rivers throughout the entire coastal range of this species, but this spawning range has contracted southward (with exception of certain Canadian rivers) coincident with human impacts in the 19th and 20th centuries. CB currently appears to be the only region providing large cohorts of striped bass to other areas. However, it is possible that more northern systems including HR have at times in the past produced year classes strong enough to drive migration to more southern systems.

Long-term migration from southern river systems is similarly unclear due to lack of long-term historical tagging data; however, there have been occasional contemporary reports of striped bass from largely nonmigratory southern stock using coastal waters to reach adjacent rivers and even undergoing long-distance migrations to northern waters (Greene et al. 2009). The degree to which striped bass originating in southern rivers migrate and successfully reproduce in more northern regions is unknown. However, it is plausible that small annual numbers of successful migrants over long timeframes may account for the observed signal of historical connectivity between SC and CB. These 2 patterns of gene flow to CB, whether strong, short-term influx from years of high reproductive success in other systems or long-term accumulation of occasional migrants, are not discernable from the present analysis, but the basic finding remains that CB has not historically been exclusively a genetic donor system to surrounding regions.

**Assignment Testing**

Although extensive characterization of the genetic composition of mixed stocks in CB and along the Atlantic coast was beyond the scope of this project, our work clearly demonstrates the suitability of microsatellite markers for assignment testing of striped bass from mixed coastal and estuarine samples. Previous studies have used RFLP of mtDNA with moderate success for this purpose; however, marked lack of sequence variation in striped bass mtDNA limits this approach, as does the strict maternal inheritance of this locus (Wirgin et al. 1990). Using the baseline panel generated in this work, 77.1% of samples from known populations could be assigned correctly under the 4-population scenario combining CB and DR. Using a more stringent criterion (80% assignment score cutoff), 65.6% of fish of unknown origin from CB and Atlantic coastal collections were successfully assigned to 1 of 4 populations (Table 4). The large majority (90.3%) of fish collected within CB demonstrated DR/CB origins, whereas a small proportion (6.4%) assigned to HR. Eighty-four percent of fish collected via NEAMAP along the Atlantic coast north of CB indicated DR/CB origins, with 1 fish assigned to NC and the balance (14.5%) to HR. Under the 5-population scenario, the proportion of fish from the NEAMAP sample assigning to HR rose to 42.1%; however, this is a proportion of only 25% of the total sample that did assign under stringent criteria. Splitting CB and DR populations in assignment analyses reduced the number of fish that could be assigned at the 80% score cutoff; therefore, assignment of fish using these baseline data appears to offer a tradeoff between ability to differentiate between these 2 populations and the number of assignable fish.

**Implications for Management**

Because of the shallow population structure in CB, it does not appear that assignment of fish to individual tributaries within this region would be feasible with the present tools.
Therefore, recruitment from individual tributaries to the migratory stock does not appear measurable, making establishment of management units with these methods within CB impracticable. The apparent flexibility in natal stream homing, in any case, would not seem to indicate support for such measures. On a coastal scale, however, genetic structuring between production centers was observed although there appears to be significant connectivity between these regions as evidenced by contemporary and historical migration analyses. The presence of genetically differentiated populations suggests potential value for regional-level management; however, the data most relevant to this issue would be relative contributions of various regions to the migratory stock, as measured by assignment testing of individuals from mixed coastal assemblages. In this scenario, under-representation of fish from a given region in coastal catches would suggest the need for consideration of regional management, or perhaps mechanisms similar to adaptive management practices currently practiced for Pacific salmonids could be implemented (Habicht et al. 2006).

Although large-scale assignment testing of individuals in the mixed migratory stock was beyond the scope of this work, we have shown clear evidence of temporally stable coastal population structure and have demonstrated that these baseline data are suitable for future determination of proportional representation of various production areas in migratory striped bass stocks. Natal origin data could further add significant information to demographic analyses of striped bass, including studies of migration patterns and differential recruitment. The striped bass stock is currently considered healthy along the Atlantic coast, and the latest stock assessment indicates it is not overfished (Atlantic States Marine Fisheries Commission 2011). A decline in stock abundance since 2004 has been observed, however, and there is evidence for increased natural mortality in CB (Jiang et al. 2007) and elsewhere along the Atlantic coast, possibly due in part to the effects of the bacterial disease mycobacteriosis (Gauthier et al. 2008; Atlantic States Marine Fisheries Commission 2011). If regional increases in natural mortality are in fact occurring, accurate modeling of their effects on overall stock health will require determination of the relative contributions of various areas to the coastal stock. For example, if disease-related natural mortality is considerably higher in CB than in other areas, this would be expected to have a large effect on the overall stock, with the magnitude depending on the relative contribution of CB to D.G. and J.C.); the Beaufort Marine Research Award in Fish Population Genetics, funded by the Irish Government under the Sea Change Programme (to J.C.).

Acknowledgments

We acknowledge the following organizations and individuals for providing samples: K. Hattala and R. Adams, New York Department of Environmental Conservation, Hudson River Fisheries Unit; T. Baum, New Jersey Bureau of Marine Fisheries; E. Durell, Maryland Department of Natural Resources; M. Fabrizio and L. Machut, Virginia Institute of Marine Science, Juvenile Finfish Survey; R. Latour and C. Bonzek, Virginia Institute of Marine Science; C. Godwin, North Carolina Department of Environmental and Natural Resources, Division of Marine Fisheries. Figure 1 was created with the assistance of Debra Gauthier.

References


Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

Saltonstall-Kennedy Grant Program, National Oceanic and Atmospheric Administration (NA08NMF4270413 to Gauthier et al. 2006).


Liu JX, Ely B. 2009. Sibship reconstruction demonstrates the extremely low effective population size of striped bass Morone saxatilis in the Santee-Cooper system, South Carolina, USA. Mol Ecol. 18:4112–4120.


Morgan RPI, Koo TSY. 1973. Electrophoretic determination of populations of the striped bass, Morone saxatilis, in the Upper Chesapeake Bay. Trans Am Fish Soc. 102:21–32.


Received February 1, 2013; First decision March 20, 2013; Accepted April 8, 2013

Corresponding Editor: Stephen Karl