Received: 29 March 2022 | Revised: 5 August 2022 | Accepted: 13 August 2022

DOI: 10.1002/tafs.10387

FEATURE ARTICLE

River-of-origin assignment of migratory Striped Bass, with implications for mixed-stock analysis

Kristopher J. Wojtusik[1] | **David L. Berlinsky**[1] | **Linas W. Kenter**[1] | **Adrienne I. Kovach**[2]

[1]University of New Hampshire, Kendall Hall, Durham, New Hampshire 03824, USA

[2]University of New Hampshire, James Hall, Durham, New Hampshire 03824, USA

Correspondence
Adrienne I. Kovach
Email: akovach@unh.edu

Abstract

Objective: The Striped Bass *Morone saxatilis* is an anadromous teleost with a native range extending north from the Gulf of Mexico into Canadian waters. Far-ranging coastal migrations support one of the most popular recreational fisheries in the United States. Identifying the underlying population genetic structure of the spawning populations and the genetic markers capable of differentiating among them advances our understanding of these economically and ecologically important fish and enables more targeted management to occur.

Methods: We used a restriction site-associated DNA sequencing approach to identify neutral and adaptive single-nucleotide polymorphisms (SNPs), and we determined the population genetic structure of 438 adult Striped Bass sampled from nine spawning locations along the Atlantic coast from the Roanoke River, United States, to the Miramichi River, Canada.

Result: The two Canadian populations (Shubenacadie and Miramichi rivers) were genetically distinct from U.S. populations and from each other. Neutral loci differentiated Striped Bass from U.S. waters into four genetically distinct populations: Roanoke River, Hudson–Kennebec River, Upper Chesapeake Bay–Potomac River–Delaware River, and Choptank River (eastern Chesapeake Bay). Outlier loci further differentiated the Delaware River from the Chesapeake Bay tributaries, suggesting that there may be local adaptation in the face of gene flow. We identified 1300 highly informative SNPs (the top 10% [with respect to the genetic differentiation index F_{ST}] of the full suite of 13,361 SNPs in our study) capable of assigning fish with at least 90% accuracy to their river of origin; through simulations, we established their applicability for conducting robust mixed-stock analyses of the coastal migratory Striped Bass fishery.

Conclusion: This study demonstrated that neutral and adaptive loci together provide evidence for fine-scale population structure of migratory Striped Bass, and these loci provide the most informative genetic panel for mixed-stock analysis of Striped Bass to date, capable of assigning fish to their spawning river of origin.

KEYWORDS

mixed-stock analysis, *Morone saxatilis*, population assignment, population genomics, population structure, Striped Bass

INTRODUCTION

The delineation of genetic stock structure is necessary for effective management of exploited fishes (Palsboll et al. 2007). Fisheries management that aligns with biological population structure aids in preserving the biocomplexity of the fishery resource, which is critical for maintaining resilience to environmental and anthropogenic pressures (Hilborn et al. 2003). Knowledge of population genetic structure is important for ensuring that the spatial scale of management matches the biological units (Reiss et al. 2009; Kerr et al. 2017), for identifying genetically compatible individuals to be used in stocking and supplementation efforts (Ward 2006), and for use in real-time genetic stock identification for the management of mixed-stock fisheries (Flannery et al. 2010; Dahle et al. 2018). Delineating genetic structure among populations that have recently diverged or have ongoing gene flow is challenging due to the high resolution needed to detect subtle genetic differentiation (Martinez et al. 2018). Prior to the genomics era, traditional genetic markers (e.g., microsatellites) sometimes lacked the resolution needed to discriminate among these subtle population differences (Hess et al. 2011).

Advances in sequencing technologies and techniques, such as restriction site-associated DNA sequencing (RADseq; Baird et al. 2008), provide the ability to randomly sample thousands of single-nucleotide polymorphisms (SNPs) distributed across an organism's entire genome. The RADseq approach and other reduced-representation sequencing approaches (Campbell et al. 2018) have become relatively commonplace in fisheries management and have proven useful in discerning subtle population structure in many marine (Benestan et al. 2015; Vendrami et al. 2017; Drinan et al. 2018; Jenkins et al. 2019) and freshwater (Chen et al. 2020) species. These sequencing advances have also been accompanied by analytical advances in the discovery and application of outlier loci (loci that yield statistically elevated population differentiation and thus are putatively under selection; Allendorf et al. 2010; Stapley et al. 2010; Gagnaire et al. 2015; Whitlock and Lotterhos 2015). Outlier loci have the potential to aid conservation efforts by identifying locally adapted populations in species of conservation concern. They also permit high-resolution differentiation of populations and provide enhanced power for population assignments at fine geographic scales (Nielsen et al. 2012; Gagnaire et al. 2015). This increased assignment accuracy has numerous applications in fisheries management, including tracking cases of illegal fishing (Martinsohn and Ogden 2009) and mixed-stock analyses of highly migratory species (Ackerman et al. 2011).

Impact Statement

Migratory Striped Bass that occur along the Atlantic coast of the USA and Canada are structured into genetically distinct populations, corresponding to their spawning river of origin. We identified a suite of genetic markers that will enable fishery managers to determine the stock composition of the mixed coastal Striped Bass fishery.

High-resolution genetic tools for population delineation and mixed-stock analysis have high applicability to the management of the Striped Bass *Morone saxatilis*, an anadromous, euryhaline, migratory teleost that is indigenous to the Atlantic and Gulf coasts of the United States and Canada (Chen et al. 2020). Within the U.S. Atlantic range, spawning stocks are comprised of geographically separate migratory and resident contingents. South of the Albemarle Sound (coastal North Carolina), stocks are largely residential, with adults spending the duration of the nonspawning season in the estuaries and coastal waters around the rivers in which they spawn. Stocks located north of the Albemarle Sound undertake an age-structured, postspawn feeding migration northward along the U.S. coastal waters or out into nearby bays in the case of Canadian populations (Waldman et al. 1990; Secor and Piccoli 2007; Rothermel et al. 2020; Secor et al. 2020). During the feeding migration, summer residency, and subsequent southerly fall migration, Striped Bass form a mixed aggregation, which supports multiple small commercial fisheries and one of the most popular recreational fisheries in the United States (NMFS 2020). Striped Bass have likely supported large and productive fisheries since North America was first colonized by indigenous peoples, and they have continued to do so since European colonizers first appeared—until populations crashed due to overexploitation in the early 1980s, leading to increased restrictions on commercial and recreational fisheries (Boreman and Austin 1985). The result of these restrictions was the recovery of the larger stocks (Chesapeake Bay and Hudson River) by the mid-1990s and the recovery of all populations by 2003 (ASMFC 2003). Spawning populations also occurred in Canadian rivers throughout New Brunswick and Nova Scotia draining into the Bay of Fundy, along the Northumberland Strait, and in the St. Lawrence River, until the late 1980s, when anthropogenic pressures (e.g., overfishing and dam building; Douglas et al. 2003; Dadswell et al. 2018) caused these populations to also decline. The closure of commercial fisheries and the implementation of regulatory restrictions to the recreational fisheries enabled the Miramichi and Shubenacadie River populations to recover naturally; subsequently, the St. Lawrence River was restored using Miramichi River-origin broodstock (Robitaille et al. 2011).

The population crash spurred research into the connectivity of migratory Striped Bass spawning stocks. In particular, the population genetic structure of these stocks has been investigated in a number of studies over the past four decades using a variety of molecular techniques. Restriction length polymorphisms (Wirgin et al. 1990), microsatellites (Robinson et al. 2004; Gauthier et al. 2013; Anderson et al. 2014; Wirgin et al. 2020), eye lens proteins (Fabrizio 1987), and SNPs (Leblanc et al. 2018, 2020) have had varying degrees of success at distinguishing spawning populations of Striped Bass. Of these studies, only three have included a comprehensive set of migratory populations in U.S. waters (Gauthier et al. 2013; LeBlanc et al. 2020; Wirgin et al. 2020) and two of those included spawning populations from Canada in addition to those from the United States (LeBlanc et al. 2020; Wirgin et al. 2020). These studies found Canadian populations to be the most distinct from one another and from U.S. populations, while in the United States they identified three regional groupings composed of the southernmost migratory rivers, including (1) the Roanoke and Cape Fear rivers, (2) the Chesapeake Bay–Delaware River complex, and (3) the Hudson and Kennebec rivers. Within the Chesapeake Bay, studies have found weak but significant east–west and north–south differentiation (Gauthier et al. 2013; LeBlanc et al. 2020; Wirgin et al. 2020). There have been conflicting results about the Delaware River, with some studies finding differentiation (Gauthier et al. 2013) and others finding no differentiation from the Chesapeake Bay (LeBlanc et al. 2020; Wirgin et al. 2020), leading LeBlanc et al. (2020) to conclude that the Chesapeake Bay–Delaware River complex functions as a metapopulation with extensive gene flow among the tributaries, with the Chesapeake–Delaware Canal as the main driver of this connectivity.

Despite the numerous studies described above, inconsistencies at a fine geographic scale—largely due to a lack of resolution in genetic markers used—warrant further study. The population structure identified in these prior studies was based solely on neutral loci. However, outlier (putatively adaptive) loci might enhance the resolution of the genetic structure and clarify the spatial scale of differentiation. Enhanced resolution would further improve the potential for characterizing the mixed fishery, which has been hampered by the low resolution of prior markers (Fabrizio 1987; Wirgin et al. 1997; Waldman et al. 2012; Gauthier et al. 2013). LeBlanc et al. (2020) assigned individuals to one of the three regions with high accuracy by using almost 1,300 SNPs, but those authors could not accurately assign individuals to the river of origin. The authors concluded that the rivers within the regions therefore were not demographically independent. Alternatively, the SNPs in their study may have lacked resolution to make finer-scale assignments. A higher-resolution panel of markers, including outlier loci, may facilitate more successful investigation of mixed-stock composition.

The recovery of Striped Bass spawning populations in the 1990s was a management success story and allowed for continued harvest by fisheries, albeit with new and more stringent regulations in place. These regulations include a complete moratorium on commercial and recreational fishing for Striped Bass in federal waters (>5 km offshore) and restricted commercial fisheries in state waters (ASMFC 1981). Striped Bass, however, still face substantial fishing pressure. A stock assessment completed in 2019 found that Striped Bass spawning stock biomass and juvenile recruitment were below threshold levels, indicating that populations were yet again in decline (NEFSC 2019). A high-resolution genetic assay capable of river-of-origin assignments would provide an important tool for management of the migratory Striped Bass stock, whereby managers can identify the fine-scale composition of mixed fisheries in different seasons and at different locations for a targeted management approach.

The objectives of this study were to (1) identify the population genetic structure across the migratory range of Striped Bass by using neutral and outlier loci, (2) perform population assignment tests to identify the finest spatial scale at which individuals can be accurately assigned, (3) identify an informative set of SNP loci to be used in future mixed-stock analyses, and (4) use simulations to test the performance of the selected loci for conducting mixed-stock analyses.

METHODS

Striped Bass DNA samples

We used DNA samples collected in previously published microsatellite and mitochondrial DNA studies of Striped Bass population structure (Wirgin et al. 1993, 2020; Robinson et al. 2004). Samples were collected from spawning adults or age-0 to age-1 juveniles from nine major spawning rivers across the migratory range of Striped Bass in U.S. and Canadian waters, including the Roanoke River; three locations within the Chesapeake Bay (Potomac River, Choptank River, and the Upper Chesapeake Bay [hereafter, "Upper Bay"]); and the Delaware, Hudson, Kennebec, Shubenacadie, and Miramichi rivers (Figure 1). We also included a collection of Shubenacadie River samples from the study of Kenter et al. (2018). These samples were obtained from individuals that were caught in the wild as juveniles and then reared to adulthood in a hatchery for aquaculture studies. Samples comprised two time periods: 1989–1998 and 2010–2016. All rivers except the Delaware River were sampled in the early time period. Three

FIGURE 1 Locations in the United States and Canada where Striped Bass were sampled in the 1990s and 2010s.

locations (Upper Bay, Hudson River, and Shubenacadie River) were sampled in both time periods, which allowed us to evaluate the temporal stability of genetic structure (see Table 1 for full sampling information). In total, we obtained 438 DNA extracts, with a minimum of 20 samples per collection (location by year). The DNA concentrations were determined using a Qubit 3.0 (Life Technologies, Inc.) and then normalized to a target concentration of 50 ng/μL for library preparation. Selected samples were those with concentrations greater than 10 ng/μL in order to have sufficient yield in library preparation.

Library preparation and sequencing

We prepared three pooled sequencing libraries for the 438 samples following the 3RADseq protocol as described by Graham et al. (2015), with one modification: we size-selected for 650–850-bp fragments on a Blue Pippin (Sage Science). The concentration of each index group was determined by using a Qubit, and the average fragment length was determined by using a TapeStation 2200 (Agilent). We calculated the molar concentration of each index group, normalized the concentrations across groups, and then pooled groups, resulting in three libraries that were submitted for sequencing at Novogene Corp. on an Illumina Hi-Seq X with PE 150 chemistry.

Filtering and single-nucleotide polymorphism calling

We used FastQC version 0.11.5 (Andrews 2010) to assess read quality before and after trimming and quality filtering. The *process_radtags* module in Stacks version 2.4 (Catchen et al. 2013) was used to demultiplex, trim reads to 140 bp (−*t*), discard reads with a Phred quality score less than 10 (−*q*), remove reads with an uncalled base (−*c*),

TABLE 1 Locations where spawning and age-0 Striped Bass were sampled, as well as collection year, references for the studies in which the samples were originally collected, gear type, and specimen life stage (Upper Bay = Upper Chesapeake Bay). Sample size indicates the number of samples included in genetic analyses.

Location	Collection year(s)	Reference	Gear type	Stage	Sample size
Miramichi River	1997, 1998	Robinson et al. (2004)	Beach seine	Age 0	31
Shubenacadie River	1997–1998, 2014	Kenter et al. (2018) and Wirgin et al. (2020)	Beach seine, wild-caught hatchery adults	Age 0	59
Kennebec River	1995	Wirgin et al. (2020)	Beach seine	Age 0	42
Hudson River	1989, 2015	Wirgin et al. (1993, 2020)	Haul seine	Adult	88
Delaware River	2010	Wirgin et al. (2020)	Electrofishing	Adult	39
Upper Bay	1989, 2016	Wirgin et al. (1993, 2020)	Gill net	Adult	81
Choptank River	1989, 1992	Wirgin et al. (1993)	Gill net	Adult	43
Potomac River	1989	Wirgin et al. (1993)	Gill net	Adult	35
Roanoke River	1989	Wirgin et al. (2020)	Angling	Adult	20
Total					438

and discard reads with adapter contamination and those failing Illumina's purity filter (*--adapter_1[_2]*, *--filter_illumina*). Reads were aligned to the Striped Bass reference genome (RefSeq accession GCF_004916995.1) using Bowtie2 version 2.4.1 (Langmead and Salzberg 2012), and we used SAMtools version 1.10 (Li et al. 2009) to remove reads with multiple alignments. Finally, we used the *gstacks* module in Stacks 2.4 to identify SNPs and genotype each individual, and the *populations* module was used to create a variant call format (VCF) file for filtering.

We developed four SNP data sets, each with different filtering criteria, to use in downstream analyses. The *populations* module or the VCFtools program (Danecek et al. 2011) was used to complete SNP filtering steps. The first data set was developed to retain the maximum number of variants for population assignment tests and SNP panel development (hereafter, the "assignment data set"). It consisted of both neutral and outlier SNPs because the latter have been shown to have high power in assigning individuals back to their population of origin (Ackerman et al. 2011; Russello et al. 2012; Jorde et al. 2018). We employed modest data filters to ensure quality control while maximizing the SNPs available for selection in the assignment panel. We set the minimum minor allele count threshold at 3 (*--min_mac*), required SNPs to be present in at least one population (*−p*), and required SNPs to be genotyped in at least 70% of individuals in a population (*−r*). To remove paralogs and null alleles, we filtered any SNP that deviated from Hardy–Weinberg equilibrium with a *P*-value less than 0.00001. Finally, we kept only one SNP per locus (*--write-single-snp*) to remove linked SNPs.

To create the next three data sets for use in characterizing population structure, we applied additional filtering to the assignment data set. First, we removed SNPs that were missing from more than 50% of individuals across the entire data set (*−R*). This resulted in our "full data set." Next, we developed a "neutral data set" to explore neutral population structure among our spawning populations. To do this, we identified and removed putatively adaptive loci from the full data set. We attempted to identify outlier loci in the full data set by using two different methods. First, we used PCAdapt (Luu et al. 2017) in R version 4.0.3 (R Core Team 2020). This approach uses a hierarchical factor model with *K* latent factors to estimate the neutral underlying population structure and to identify loci that are statistical outliers in terms of the strength of their association with this structure. We determined the optimum *K*-value to retain for the analysis by considering both the scree plot and the principal components analysis (PCA) plots produced by PCAdapt. An optimum *K*-value of 6 was chosen because at this value on the scree plot, the eigenvalues stopped corresponding to population structure and there was no apparent population structure in the PCA plots. To control for false discoveries, *P*-values were transformed into *Q*-values by using the R package *Q*-value (Storey et al. 2022). Loci with *Q*-values of 0.05 or less were assumed to be significant outliers. We also used OutFLANK (Whitlock and Lotterhos 2015) to identify potential outliers. OutFLANK estimates the distribution of genetic differentiation index F_{ST} values at neutral loci by fitting the data to a chi-square distribution after trimming excessively high and low F_{ST} values, as these loci may be under selection. The empirical untrimmed data are then compared to the chi-square distribution, and outliers are identified as those outside the expected distribution. We thinned our data set to 1 SNP per 10-kb window and used the remaining SNPs to obtain the chi-square distribution. Again, any loci with a *Q*-value of 0.05 or less were considered significant outliers. We removed identified outliers to create a putatively neutral SNP data set (i.e., the neutral data set) and retained the loci identified as outliers to create the "outlier data set."

We also explored the effects of missing data on our full data set given the relatively modest missing data filter (50%). To do this, we followed the iterative filtering approach of O'Leary et al. (2018), which involved the stepwise removal of individuals with missing data exceeding thresholds of 90–30%, alternating with removal of loci with missing data exceeding thresholds of 60–90%. This resulted in a stringently filtered data set comprising loci with no more than 10% missing genotypes and individuals with no more than 30% missing data, which we used for quality control validation of the population structure analyses conducted on the full data set (see below).

Genetic diversity

The neutral data set was used to derive metrics of genetic diversity for Striped Bass sample collections. We used Genodive (Meirmans 2020) to calculate expected heterozygosity (H_e), observed heterozygosity (H_o), and the inbreeding coefficient (G_{is}). To avoid bias from having related individuals in the data set, we used the *relatedness2* function in VCFtools to identify full-sibling pairs identified with a probability of 0.25. We identified five possible full-sibling pairs (four in the Shubenacadie River and one in the Choptank River); one individual from each pair was removed from all data sets.

Population structure

We used the full, neutral, and outlier data sets in the population genetic structure analyses as follows. We calculated pairwise F_{ST} among sampling rivers and conducted significance testing in Genodive with 10,000 permutations; we corrected for multiple tests by using Myriads (Carvajal-Rodríguez 2018). To assess the genetic clustering patterns among individuals, we used the R packages Adegenet (Jombart 2008) and Ade4 (Dray and Dufour 2007) to perform an individual-based PCA, and ggplot2 was used to visualize the results. We also used Adegenet to perform a discriminant analysis of principal components (DAPC) to evaluate genetic differentiation among sampling locations and to compare the clustering patterns provided by the neutral and outlier data sets. To evaluate potential impacts of missing data on our analyses of population structure, we also performed PCA and DAPC on the stringently filtered data set (i.e., the data set generated with the O'Leary et al. 2018 filtering criteria).

We also assessed population structure for the neutral data set by using the Bayesian clustering algorithm, STRUCTURE version 2.3.4 (Pritchard et al. 2000). We performed 10 iterations for K-values of 1–10, with a burn-in length of 10,000 and a run length of 100,000 Markov chain–Monte Carlo generations. We employed the admixture model with correlated allele frequencies and the locprior model because this model is robust to weak population differentiation, thus providing higher-resolution population structure, and is unbiased to unbalanced sample sizes (Hubisz et al. 2009). For more direct comparison with earlier studies that did not use the locprior model, we also ran STRUCTURE without sampling locations as prior information. The best value of K was determined from the plateau in values of $\ln(P[D])$ (Pritchard et al. 2000) and the ΔK method (Evanno et al. 2005) implemented using STRUCTURE HARVESTER (Earl and vonHoldt 2012), as well as by examination of the bar plots produced using Clumpak (Kopelman et al. 2015).

We tested for temporal stability of population structure by conducting an analysis of molecular variance (AMOVA) using the neutral data set in the Pegas (Paradis 2010) AMOVA implementation within the R package Poppr (Kamvar et al. 2014) for the locations that were sampled during both time periods: Upper Bay, Hudson River, and Shubenacadie River. We assessed isolation by distance on two population groupings: all locations and only U.S. locations. To do this, we performed a Mantel test with matrices of genetic distance and geographic distance among pairs of spawning rivers using the R package Adegenet. Pairwise geographic distances to the mouths of rivers were measured along the coast. In the case of the Upper Bay location and the Delaware River, it was assumed that Striped Bass use the Chesapeake–Delaware Canal, and we measured around the coast for Long Island, New York, and Cape Cod, Massachusetts. Genetic distances were calculated using the R package Hierfstat. We then used the MASS package (Venables and Ripley 2002) in R to visualize the results with a two-dimensional density estimation to discern whether the resulting pattern was due to consistent spatial genetic differentiation or was attributable to distant and differentiated populations.

Population assignment

We used the assignment data set to (1) assess the power of the data to correctly assign individuals to their population of origin, (2) determine the finest scale of structure at which accurate assignments could be made, and (3) identify the most informative SNPs in the data set to create a genetic panel for use in future genetic stock identification analyses. We did so by using two approaches.

The first approach, AssignPOP (Chen et al. 2018), uses a supervised machine-learning framework to implement a Monte Carlo cross-validation procedure and PCA using training and test data sets that are independent of each other. AssignPOP allows users to test varying proportions of individuals from each population to be used in the training data set, thus allowing users to determine whether training and test sample sizes lead to bias in assignment results. To this end, we set the function *train.inds* to 0.5, 0.7, and 0.9 to use 50, 70, or 90% of the individuals from each population in the training set. The second approach, the R package Rubias (Moran and Anderson 2019), employs Bayesian inference from a conditional stock identification model and uses the leave-one-out cross-validation method that permits stock identification accuracy while reducing bias in reporting unit proportions (Anderson et al. 2008). Assignment accuracy from both AssignPOP and Rubias was used to determine the accuracy of the assignment data set in assigning individuals back to their population of origin. Individuals with an assignment accuracy (AssignPOP) or a posterior probability (Rubias) of 80% or higher were considered correctly assigned to the population.

We used the *train.loci* parameter in AssignPOP to estimate the minimum number of markers needed for an accurate assignment of the training set. Finally, the *check.loci* function was used to create a list of the top 10% of SNP loci with respect to F_{ST}; we thinned our assignment data set to include only those most polymorphic loci and again tested the assignment accuracy of the loci using AssignPOP.

Mixed-stock simulations

We used Rubias to run mixed-stock simulations, and all individuals from the spawning locations sampled at the 1,300 highest-F_{ST} loci were used as the genetic baseline. The *assess_reference_loo()* function in Rubias carries out simulations of mixtures by using the leave-one-out approach of Anderson et al. (2008), and we used this function to test (1) the power to assign unknown individuals from a mixture sample to rivers of origin, (2) the influence of mixture size on assignment accuracies, (3) the influence of mixture proportions on assignment accuracies, (4) whether the number of individuals in the reference data set influenced the assignment results, and (5) whether assignment power increased when admixed rivers were grouped into a larger reporting unit. To do this, we first considered each spawning location as a separate reporting unit. We then varied mixture size using 50, 100, and 200 individuals, and we tested two different mixing proportion sets in which we varied the mixing proportions of the two populations that are likely to contribute the most individuals to the mixed stock: the Chesapeake Bay tributaries and the Hudson River. Individuals were randomly removed from each reference river until all locations had 20 individuals; this was done to determine whether the sample size of the reference data set had an influence on assignments. Finally, we grouped the Potomac River, Choptank River, Upper Bay, and Delaware River into one reporting group, as had been done in previous studies (Gauthier et al. 2013; LeBlanc et al. 2020; Wirgin et al. 2020), to determine whether this would increase the accuracy of assignments to this region (Table 2). All simulations were run with 100 repetitions and used the "resample-over-gene-copies" resampling method (i.e., the "CV–GC" method of Anderson et al. 2008). We removed the Kennebec River from the simulations because of its genetic similarity with the Hudson River (Rubias could not distinguish between the two rivers) and because it is not likely to contribute many individuals to the mixed stock.

RESULTS

Single-nucleotide polymorphism filtering

We obtained 652 million raw paired-end reads, with an average of 1.4 million reads/individual. Stacks initially called 80,330 SNPs; after quality control and filtering, the assignment data set contained 13,361 SNPs and the full data set (after removing individuals with excessive missing data) contained 9,492 SNPs. While OutFlank did not identify any outliers in the full data set, PCAdapt identified 140 outlier loci (i.e., the outlier data set), which were removed from the full data set to yield the 9,352 SNPs of the neutral data set (Table 3). The stringent filtering (10% missing genotypes and individuals with no more than 30% missing data) resulted in 4,275 SNPs. The average depths for individuals and loci were 24× and 26×, respectively, for the assignment data set and 26× and 30×, respectively, for the full data set.

Genetic diversity

Measures of genetic diversity were largely consistent across sites, albeit with slightly lower heterozygosity values in the Hudson, Kennebec, Shubenacadie, and Miramichi rivers than in the other rivers. Overall, H_e and H_o ranged from 0.04 to 0.15 across spawning locations (Table 4). The G_{is} ranged from −0.033 to 0.014 and was negative for all but the Roanoke River site (Table 4).

TABLE 2 Mixture simulation trials per Striped Bass spawning location (Upper Bay = Upper Chesapeake Bay). Column headings indicate the reason for running the trial, while bolded numbers and names indicate what was changed from one trial to the next. Except for trial 5, the number of individuals in the reference data set per river was as follows: Roanoke River ($n = 20$), Potomac River ($n = 35$), Choptank River ($n = 43$), Upper Bay ($n = 81$), Delaware River ($n = 39$), Hudson River ($n = 88$), Shubenacadie River ($n = 59$), and Miramichi River ($n = 31$). “Ches_Del” is the abbreviation used for the combined reporting unit that included all Chesapeake Bay tributaries and the Delaware River.

						Trial 5: grouping rivers into reporting units	
Variable or location	**Trial 1: power of assignment**	**Trial 2A: small mixture sample size**	**Trial 2B: large mixture sample size**	**Trial 3: increased Hudson River contribution to mixture**	**Trial 4: reduced number of individuals in reference data set (20 per location)**	**Reporting unit**	**Mixture size or proportion**
Mixture size	100	**50**	**200**	100	100		100
Mixing proportion							
Miramichi	2.5	2.5	2.5	2.5	2.5	**Miramichi**	**2.5**
Shubenacadie	2.5	2.5	2.5	2.5	2.5	**Shubenacadie**	**2.5**
Hudson	30	30	30	**50**	30	**Hudson**	**30**
Delaware	10	10	10	10	10	**Ches_Del**	**60**
Upper Bay	20	20	20	10	20		
Choptank	15	15	15	10	15		
Potomac	15	15	15	10	15		
Roanoke	5	5	5	5	5	**Roanoke**	**5**

Population structure

Pairwise F_{ST} values for the neutral data set ranged from 0.000 to 0.151 across spawning location pairs (Table 5). The highest values were between the two Canadian collections (Shubenacadie and Miramichi rivers) and the U.S. collections, and the lowest values were among the rivers of the Chesapeake Bay and the Delaware River (Table 5). Although some F_{ST} values were very small, all comparisons were statistically significant, likely due to the large number of loci used, with the exception of the Delaware River and the Upper Bay collections ($F_{ST} = 0$). Pairwise F_{ST} values for the outlier and full data sets showed patterns similar to those for the neutral data set, with the lowest values found among the Chesapeake Bay–Delaware River system (0.00–0.01) and the highest values among the Canadian and U.S. collections (Tables S1 and S2 in the Supplemental Material available in the online version of this article). Correlations of geographic and genetic distance were significant when all sampling locations were included ($r = 0.87$, $P < 0.005$), but when isolation by distance was assessed on U.S. locations only, there was no significant pattern (Figure S1 in the Supplemental Material available in the online version of this article).

TABLE 3 Results of the filters used sequentially to create the final single-nucleotide polymorphism (SNP) data sets for Striped Bass.

Filter	Number of SNPs
All SNPs identified	80,330
Minor allele count (minimum MAC threshold = 3)	34,226
SNPs in 1 population and 70% of individuals	15,329
Hardy–Weinberg equilibrium	13,361
Single SNP per locus	13,361 (assignment data set)
SNPs missing from 50% of individuals	9,492 (full data set)
Outliers identified in PCAdapt	140 (outlier data set)
Outliers removed	9,352 (neutral data set)

TABLE 4 Observed heterozygosity (H_o), expected heterozygosity (H_e), and inbreeding coefficient (G_{is}) of Striped Bass sampled at spawning locations (Upper Bay = Upper Chesapeake Bay).

Location	H_o	H_e	G_{is}
Miramichi	0.086	0.084	−0.033
Shubenacadie	0.113	0.111	−0.019
Kennebec	0.043	0.042	−0.023
Hudson	0.098	0.095	−0.03
Delaware	0.145	0.144	−0.011
Upper Bay	0.152	0.145	−0.052
Choptank	0.144	0.141	−0.025
Potomac	0.137	0.136	−0.009
Roanoke	0.147	0.149	0.014

The AMOVA found no significant differences for the three paired temporal replicates ($\text{phi}_{st} = 0.0001$, $P = 0.96$), suggesting temporal stability in the population structure across the sampling years for Upper Bay, Hudson River, and Shubenacadie River. The phi_{st} among populations was two orders of magnitude larger than that among temporal replicates, although among-population differences were only marginally significant for these three rivers ($\text{phi}_{st} = 0.041$, $P = 0.0609$). We note here that although Pegas reports these results as phi_{st}, with biallelic data they are equivalent to F_{ST} (Meirmans and Liu 2018).

In a PCA with the full data set, the first three principal components explained a total of 10.6% of the variation seen in the data. The Canadian locations formed two separate clusters, and all of the U.S. locations were grouped together and formed a third cluster (Figure 2A). When only U.S. locations were included in the analysis, the Roanoke

TABLE 5 Pairwise genetic differentiation index F_{ST} values, calculated using the neutral data set, for Striped Bass sampled in spawning rivers (Upper Bay = Upper Chesapeake Bay). Values with an asterisk are significant ($P < 0.05$).

Location	Miramichi	Shubenacadie	Kennebec	Hudson	Delaware	Upper Bay	Choptank	Potomac
Miramichi	–							
Shubenacadie	0.149*	–						
Kennebec	0.103*	0.060*	–					
Hudson	0.121*	0.100*	0.005*	–				
Delaware	0.133*	0.125*	0.009*	0.012*	–			
Upper Bay	0.122*	0.118*	0.007*	0.012*	0.000	–		
Choptank	0.140*	0.133*	0.013*	0.021*	0.004*	0.005*	–	
Potomac	0.136*	0.124*	0.010*	0.011*	0.001*	0.002*	0.010*	–
Roanoke	0.151*	0.137*	0.046*	0.031*	0.025*	0.026*	0.036*	0.026*

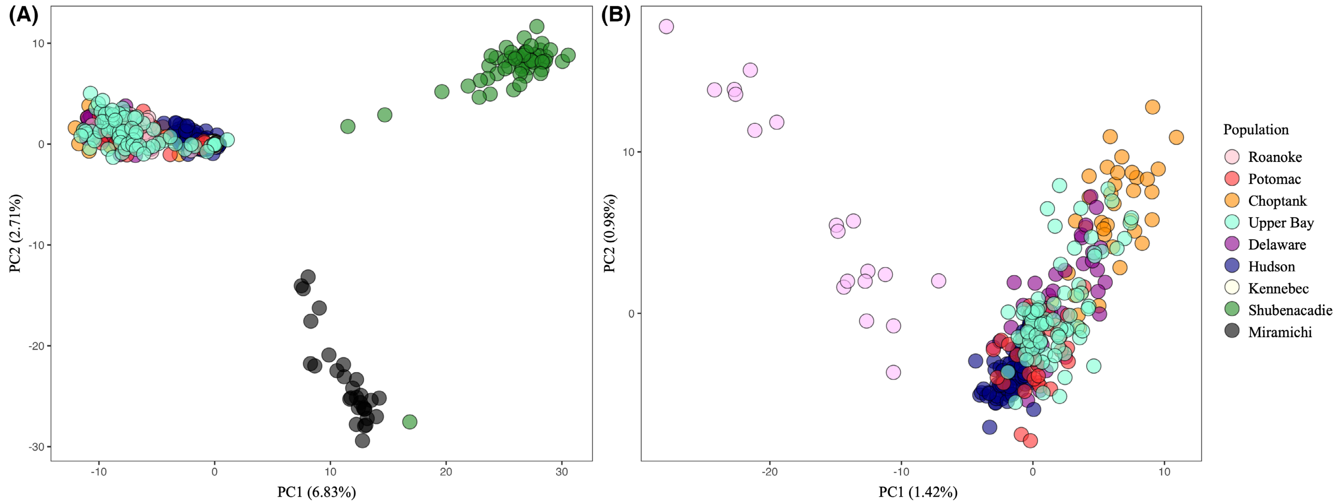

FIGURE 2 Principal components analysis plots of Striped Bass samples collected at nine spawning locations (rivers) using the full data set: (A) all spawning locations and (B) only U.S. spawning locations (PC = principal component). Dots represent individual samples, and colors correspond to the sampling location.

River clustered separately from the other U.S. locations (Figure 2B). The clustering pattern obtained when using the neutral data set was similar to that observed with the full data set (Figure S1). Similarly, the results from using the outlier data set were largely the same as those from the full data set (Figure S2). A PCA conducted with the stringently filtered data set of 4,275 SNPs recovered the same clustering pattern, suggesting that there were no impacts of missing data on the observed patterns of population structure (Figure S3).

We used the full, neutral, and outlier data sets to explore population structure in the DAPC using spawning locations as a priori groups. The DAPC clustering patterns using the full data set (i.e., the combination of all neutral and outlier loci) with a priori population groupings were largely similar to the pattern obtained with the neutral data set (Figure S4). Therefore, results from only the neutral and outlier data sets are reported here. Using the neutral data set, DAPC showed three distinct clusters, comprised of the two Canadian locations (Miramichi and Shubenacadie rivers) separately and all of the U.S. spawning locations together (Figure 3A). When only U.S. locations were included in the analysis, DAPC again showed three distinct clusters: the Roanoke and Hudson rivers each clustered separately, and the Chesapeake Bay locations and the Delaware River clustered together (Figure 3B). The DAPC of the stringently filtered data set of 4,275 SNPs recovered the same patterns of population structure as observed with the full and neutral data sets, again suggesting that there were no artifacts of missing data in our population structure analysis (Figure S5). The DAPC clustering patterns obtained using the outlier data set showed greater separation of the Roanoke and Delaware rivers from each other and from other locations as well as some separation of the Chesapeake Bay tributaries (Potomac River, Choptank River, and Upper Bay; Figure 3C). Significant separation of the Roanoke River and a few of the Chesapeake Bay tributaries can be seen when locations are plotted with loading 3 (Figure S6).

For the STRUCTURE analysis using the locprior model, both Δ*K* and ln(*P*[*D*]) suggested a *K*-value of 6 (Figure S7A,B). The six clusters were as follows: (1) Roanoke River; (2) Potomac River; (3) Choptank River, Upper Bay, and Delaware River; (4) Hudson and Kennebec rivers; (5) Shubenacadie River; and (6) Miramichi River (Figure 4). Analysis without sampling location as prior information yielded similar results; although Δ*K* and ln(*P*[*D*]) suggested a *K* of 5 (Figure S8), the bar plots were most stable at a *K* of 6, for which they showed the same clustering pattern as with the locprior model (Figure S9).

Population assignment

Population assignment analyses using the assignment data set showed high self-assignment of individuals back to their river of origin, with largely similar results from Rubias and AssignPOP (Table 6). Assignment rates ranged from 90% to 100% with both methods, except for the Kennebec River, which had an average assignment of 42% in Rubias and 96% in AssignPOP. The majority of misassigned individuals in the Kennebec River were assigned to the Hudson River (Table S3).

Assignment accuracy was 90% or better for all populations and all proportions of individuals tested (50, 70, and 90%; Table S4). There was no apparent bias in

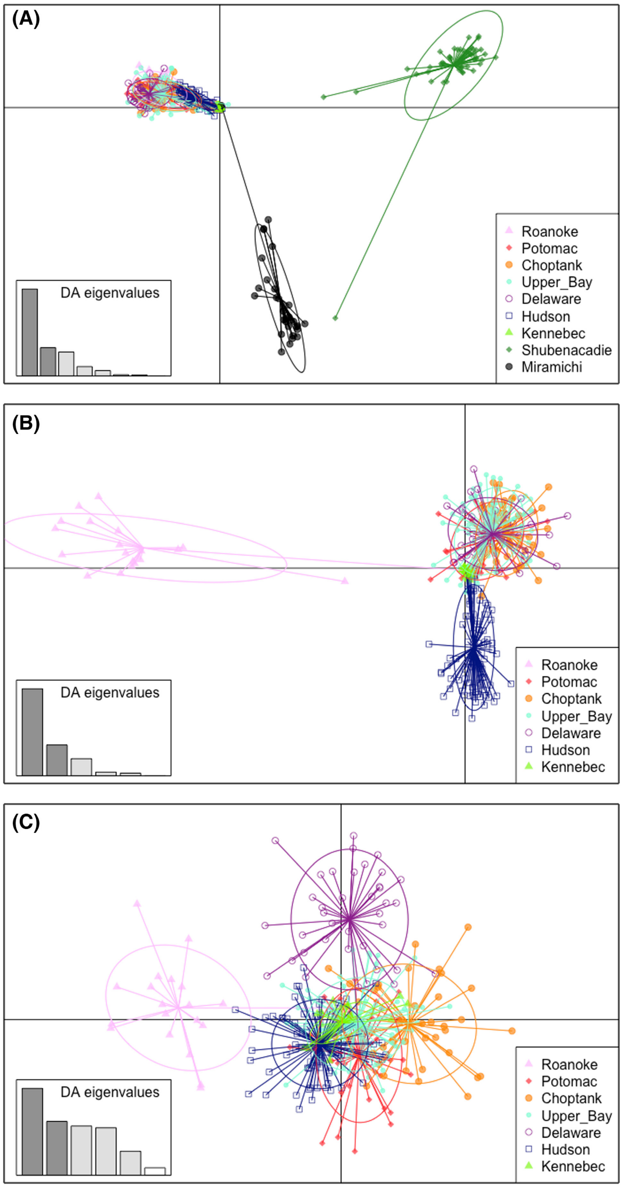

FIGURE 3 Discriminant analysis (DA) of principal components plots of Striped Bass from nine spawning locations in the United States and Canada using 9352 neutral single-nucleotide polymorphisms (SNPs) and 140 outlier SNPs: (A) all nine U.S. and Canadian locations examined using neutral SNPs; (B) U.S. spawning locations examined using neutral SNPs; and (C) U.S. spawning locations examined using outlier SNPs.

sample size, so we report results using the 70% proportion of individuals for visualization. Assignment accuracy was similar for all proportions (10–100%) of loci used, across 30 iterations, with mean accuracies of 91–97% (Table S4). To identify a panel of the most informative SNPs, we conducted further assignment tests with the top 1.0, 2.5, 5.0, and 10.0% of loci based on the highest F_{ST} values. We found that assignment accuracy using the top 1.0% and 2.5% of loci was variable across populations (0–98% accurate; Figure 5). Accuracies for the top 5% and 10% of loci were largely consistent and high (88–100%; Figure 5). All populations had an assignment accuracy >90% using the top 10% of loci with respect to F_{ST}. We identified these high-resolution SNPs as an "assignment panel" and evaluated their performance in mixed-stock simulations.

Mixed-stock simulations

Mixture simulations using the assignment panel showed high assignment to the river of origin (Figure S10), and median accuracies were 92–100% for all trials with full reference population sample sizes (Table 7). Reduction in the number of individuals in the reference data reduced the accuracies for the Hudson River, Potomac River, and Upper Bay to 36, 8, and 1%, respectively (Table 7). Grouping of the reference locations for the Chesapeake Bay system and the Delaware River into a single reporting unit reduced the number of misassignments for that system and increased the median assignment to 100% for every location (Figure S11).

DISCUSSION

Delineating the genetic stock structure of anadromous species in the face of gene flow (due to straying) can be challenging (McLean and Taylor 2001), and it is now more feasible due to modern sequencing technologies and associated genomic tools (Sutherland et al. 2021). Here, we identified 13,361 SNP loci from 3RADseq and developed multiple data sets consisting of neutral loci, outlier loci, and a combination of the two types of loci to explore the population genetic structure of Striped Bass within their migratory range in U.S. and Canadian waters. Neutral loci confirmed patterns of population structure identified in prior studies (LeBlanc et al. 2020; Wirgin et al. 2020), while outlier loci identified finer-scale genetic differences than were previously found. A panel of 1,300 discriminatory SNPs (both neutral and adaptive) provided high-resolution assignment (≥89%) of Striped Bass to their river of origin—a higher resolution than has been possible to date. These findings and genetic resources will facilitate fine-scale management of the coastal mixed fishery for Striped Bass in U.S. waters.

Population structure

From our analysis of neutral and adaptive variation, we found evidence for differentiation of U.S. and Canadian

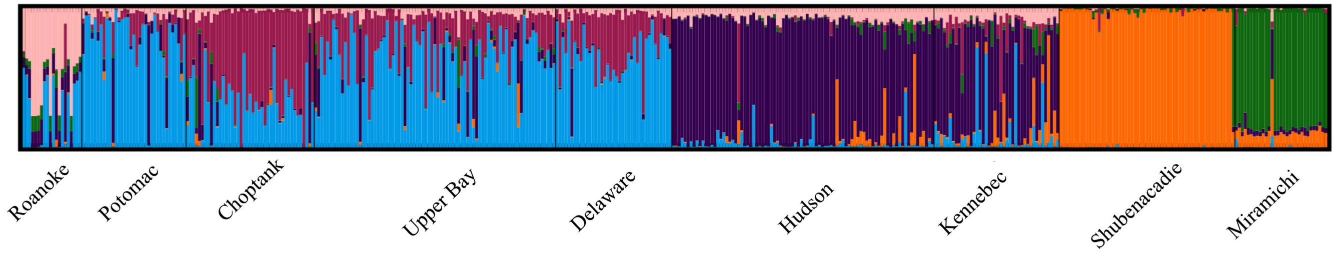

FIGURE 4 STRUCTURE results for nine Striped Bass spawning locations based on $K = 6$ population clusters. Each vertical bar represents an individual sample, and the different colors represent the contribution of each of the K genetic clusters to each sample's genotype.

TABLE 6 Self-assignment results from Rubias and AssignPOP for Striped Bass sampled from nine spawning rivers (Upper Bay = Upper Chesapeake Bay).

Location	Rubias	AssignPOP
Miramichi	0.96	0.96
Shubenacadie	0.98	0.96
Kennebec	0.42	0.96
Hudson	1.00	0.97
Delaware	1.00	0.97
Upper Bay	0.97	0.97
Choptank	0.90	0.89
Potomac	0.94	0.92
Roanoke	1.00	0.92

migratory Striped Bass spawning populations on a fine scale. Neutral loci distinguished four major spawning areas in U.S. waters: the Roanoke River, Hudson River, and eastern (Choptank River) and western (Potomac River–Upper Bay) portions of the Chesapeake Bay. Striped Bass from the Delaware River were found to be genetically similar at neutral loci to fish from the western portion of the Chesapeake Bay (Potomac River and Upper Bay), suggesting gene flow between these regions; Striped Bass from the Kennebec River were genetically similar to those from the Hudson River. The latter finding is likely due to a stocking program implemented by the state of Maine from 1982 to 1991 that introduced juvenile Striped Bass of Hudson River origin to the Kennebec–Androscoggin River system (Flagg and Squires 1994; LeBlanc et al. 2020; Wirgin et al. 2020). The two Canadian locations, the Shubenacadie and Miramichi rivers, were strongly differentiated from each other ($F_{ST} = 0.149$) and from all U.S. spawning locations ($F_{ST} = 0.060–0.151$), with little to no gene flow between them and seemingly none with U.S. populations. Differentiation among U.S. locations was much lower ($F_{ST} = 0.000–0.046$), with the strongest differentiation observed between the Roanoke River and the other populations ($F_{ST} = 0.025–0.046$). Despite being geographically proximal to the Chesapeake Bay, the Roanoke River has been shown to be one of the most distinct U.S. Striped Bass populations by our study and previous studies (LeBlanc et al. 2020; Wirgin et al. 2020). This may be due to the geographic barrier posed by the Outer Banks of North Carolina, which likely minimizes the movements of Roanoke River-spawning Striped Bass beyond the Albemarle Sound as well as minimizing the straying of Striped Bass from other areas into the Roanoke River.

Population differentiation followed a pattern of isolation by distance across the full migratory range, including Canadian locations. However, genetic differentiation was not correlated with geographic distance when only the U.S. locations were considered. This suggests that the differentiation of the two Canadian rivers drives the isolation by distance pattern and that the differentiation within U.S. waters is on a finer scale but varies spatially. For example, genetic similarity between the Hudson and Kennebec rivers due to legacy stocking is on a larger geographic scale than the distance that separates the genetically distinct areas within the Chesapeake Bay.

Previous studies identified three genetically distinct regional groups of migratory U.S. Striped Bass: Roanoke River and North Carolina, Hudson–Kennebec River, and a Chesapeake Bay–Delaware River complex that functions as a metapopulation (LeBlanc et al. 2020; Wirgin et al. 2020). Our results corroborate these findings and also suggest finer-scale structure within the Chesapeake Bay–Delaware River complex. Specifically, we found that the Choptank River on the eastern shore of the Chesapeake Bay was discrete from all other sampled populations within the bay and from the Delaware River population. This east–west differentiation is consistent with patterns found by Gauthier et al. (2013), whereby the Potomac River and Upper Bay region were different from the three southern locations (the Rappahannock, York, and James rivers). While those authors did not sample any rivers on the east side of the bay, their finding of differentiation within the bay—specifically, the Potomac

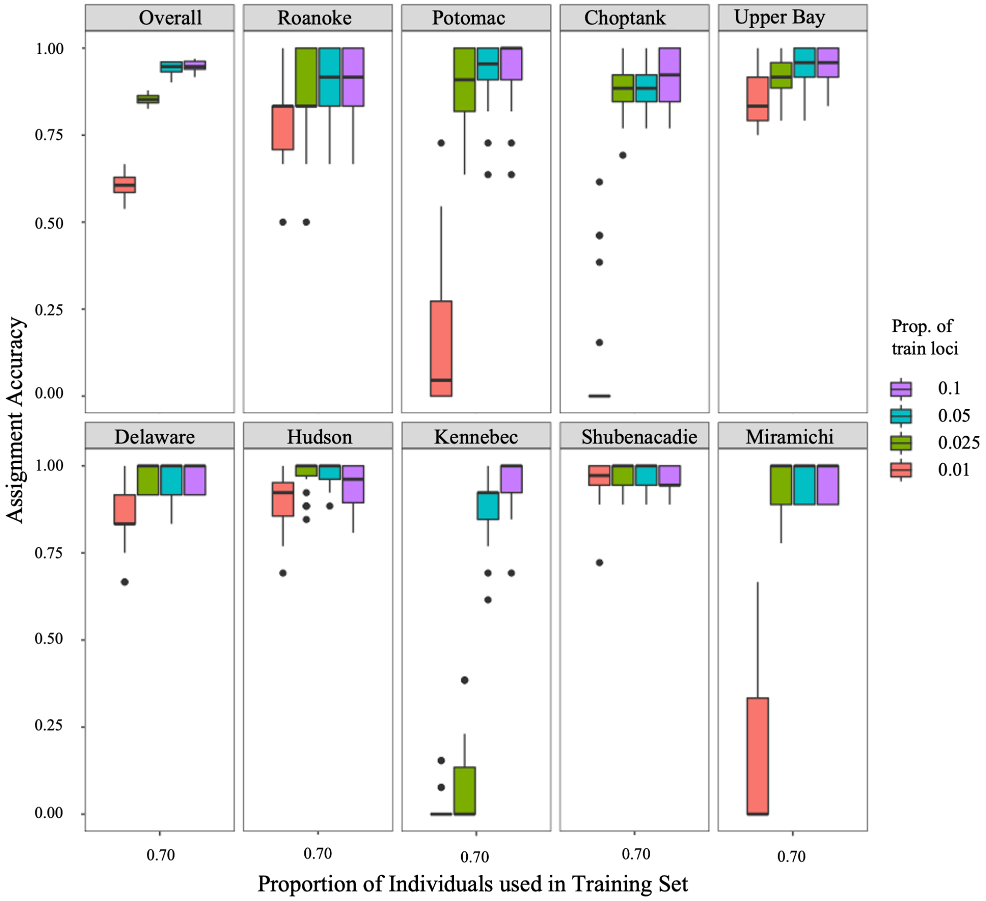

FIGURE 5 Assignment accuracy from nine Striped Bass spawning locations using subsets of 13,361 single-nucleotide polymorphisms (SNPs) in AssignPOP. Results are shown for the top 10.0, 5.0, 2.5, and 1.0% of loci based on the highest F_{ST} values in the data set. Colors represent the different proportions of loci used in the analysis, and box plots portray the median (thick black line), interquartile range (ends of boxes), and outliers (black dots).

TABLE 7 Mixture simulation median assignment (%) results per Striped Bass spawning location (Upper Bay = Upper Chesapeake Bay). Trials are defined in Table 2. Mean assignments (%) are given in parentheses. "Ches_Del" is the abbreviation used for the combined reporting unit that included all Chesapeake Bay tributaries and the Delaware River.

						Trial 5	
Location	**Trial 1**	**Trial 2A**	**Trial 2B**	**Trial 3**	**Trial 4**	**Reporting unit**	**Assignment (%)**
Miramichi	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	Miramichi	100 (100)
Shubenacadie	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	Shubenacadie	100 (100)
Hudson	100 (98)	100 (98)	100 (98)	100 (98)	36 (47)	Hudson	99 (97)
Delaware	100 (99)	100 (99)	100 (100)	100 (100)	100 (100)	Ches_Del	99 (96)
Upper Bay	99 (89)	100 (88)	100 (88)	100 (88)	1 (29)		
Choptank	99 (81)	99 (81)	100 (82)	100 (81)	99 (81)		
Potomac	95 (67)	92 (65)	100 (68)	100 (67)	8 (37)		
Roanoke	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	Roanoke	100 (100)

River–Upper Bay similarity—matches our findings and provides further evidence for fine-scale structuring within the bay. Additionally, outlier loci in our study further differentiated Striped Bass that spawn in the Upper Bay and Potomac River from those spawning in the Delaware River, suggesting a potential role for local adaptation at the level of individual rivers. As our study only included three tributaries of the Chesapeake Bay, future research that includes samples from additional tributaries and uses our high-resolution markers may be warranted to further understand the population substructure within the Chesapeake Bay.

Adaptive divergence has been shown to exist in species despite geographically proximal populations and high levels of gene flow among populations (Nielsen et al. 2009). It has also been shown that contemporary gene flow does not override historical isolation with respect to population structure in highly vagile species (Avise et al. 1987; Bermingham et al. 1992; Schneider et al. 1998). Therefore, it is possible that adaptive differences may persist in the face of contemporary gene flow between the Delaware River and the Chesapeake Bay. Additionally, previous studies using neutral loci found small but significant differences between the Delaware River and the Chesapeake Bay (Waldman and Wirgin 1995; Bielawski and Pumo 1997; Gauthier et al. 2013), suggesting that gene flow is modest. This idea is reinforced by the results from our assignment tests, in which assignment accuracies to the Delaware River and Chesapeake Bay tributaries were high.

Findings from our analysis of neutral loci are in agreement with those of two recent studies using microsatellite loci (Wirgin et al. 2020) and SNPs (LeBlanc et al. 2020) with regard to the strong differentiation of Canadian Striped Bass. The greater differentiation of the Striped Bass in Canadian rivers compared to those in U.S. waters may be due to differences in migratory patterns in Canadian and U.S. waters. Striped Bass in U.S. waters undertake substantial north–south coastal migrations over greater distances than Striped Bass in Canadian waters, thereby encountering more opportunities for straying among rivers. Striped Bass in Canadian waters undertake short migrations to larger bodies of water (Bay of Fundy and Gulf of St. Lawrence) that are proximal to the river in which they reproduce (LeBlanc et al. 2020), thus limiting opportunities for straying. Specifically, Canadian Striped Bass from western Nova Scotia and eastern New Brunswick occupy areas throughout the Bay of Fundy after spawning occurs (Rulifson et al. 2008) and then overwinter in warmer coastal waters and estuaries around their natal rivers; however, there is very little north–south movement (Rulifson and Dadswell 1995). Similarly, while Striped Bass from the Miramichi River in New Brunswick have been seen as far as the Labrador coast (Andrews et al. 2019), there is no indication that fish from these rivers have moved south along the eastern coastline of Nova Scotia, and overwintering habitats occur in and around the river (Douglas et al. 2009). There have been examples of Canadian fish being captured as far south as Virginia and Hudson River fish being captured in the Bay of Fundy (Waldman et al. 1990; Rulifson et al. 2008), but such occurrences are rare and those examples occurred during the nonbreeding season. Thus, the shorter migratory distances of Canadian Striped Bass compared to those in U.S. waters result in much less straying and minimize the contribution of Canadian rivers to U.S. Striped Bass populations.

Applications for characterizing the mixed coastal fisheries

Although much has been learned over the last few decades about the composition of the mixed coastal U.S. Striped Bass fishery, there are many remaining unknowns, particularly with respect to the fine-scale (river) composition of mixed aggregations in specific locations and seasons and across years. Although it is of value to managers, this finer-scale information has been challenging to obtain due to the limitations of accurate river-of-origin assignments, yearly variation in stock composition along the Atlantic coast (Wirgin et al. 1993), and the long-distance migrations undertaken by Striped Bass (Callihan et al. 2014, 2015). The first mixed-stock analyses based on morphometrics found that Chesapeake Bay-origin fish comprised the majority of fish caught in the mixed fisheries from Maine to North Carolina (Berggren and Lieberman 1978). Subsequent genetic mixed-stock analyses conducted in the late 1980s and 1990s on collections from Rhode Island and New York found that the Hudson River contribution to the fishery was nearly equal to or greater than the Chesapeake Bay contribution (Fabrizio 1987; Wirgin et al. 1993, 1997). Most recently, analysis of collections from New Jersey, Delaware Bay, and North Carolina found that the Chesapeake Bay was again the largest contributor to the Striped Bass mixed fishery, which was credited to the recovery of the Chesapeake Bay stocks (Waldman et al. 2012). These studies highlighted the contribution of the two largest populations, the Chesapeake Bay and the Hudson River, to the Striped Bass mixed-stock fishery but were limited in resolution by their genetic markers and by the limited sampling locations for each regional fishery. Accurate and high-resolution characterization of the coastal mixed fishery is of high relevance to managers given the current population declines of Striped Bass. Determining the contribution of individual rivers to the mixed fishery would allow for more targeted management of the fishery (i.e., with spatial and temporal resolution) and would minimize the chances of a single spawning river being disproportionately harvested.

This is the first study to identify genetic markers with high resolution to assign Striped Bass individuals to their river of origin. Previous studies attempting to assign Striped Bass to a river of origin were met with limited success due to limited resolution of the genetic markers. Using 14 microsatellite loci, Gauthier et al. (2013) were able to assign 60% of unknown individuals to one of three regional groups: the Hudson River, Chesapeake Bay–Delaware River, and North Carolina. Wirgin et al. (2020), using a panel of eight microsatellite loci, met with slightly better success, reporting self-assignment rates of 65–74% for the same regional groupings. Using 1,256 neutral SNPs, LeBlanc et al. (2020) assigned 99% of Striped Bass with more than 80% confidence to the correct regional groupings but had only 53% correct assignment to river of origin. In our study, we used the top 1,300 polymorphic SNPs from our full data set of 13,361 SNPs to assign individuals to their river of origin with 89–97% accuracy and to the three regions with 100% accuracy.

Low genetic differentiation among rivers can lead to misassignments and may indicate that rivers should be aggregated together into reporting groups. A few of the misassignments were within the Chesapeake Bay–Delaware River complex: the Choptank and Potomac rivers had four and two misassignments, respectively, to the Delaware River; and the Upper Bay had one misassignment to each of the Delaware and Shubenacadie rivers. The latter result is surprising given the high level of differentiation between Striped Bass in U.S. and Canadian waters, and it may represent a rare migrant or a sample labeling error. The low level of misassignments suggests that analyses at the level of individual rivers are warranted. The Kennebec River had the most misassignments at 24. Seventeen of those misassignments were to the Hudson River, while seven were to the Upper Bay. These results were similar to those of LeBlanc et al. (2020), who grouped the Kennebec River with the Hudson River. The poor assignment results were obtained using Rubias, whereas AssignPOP had much fewer misassignments for the Kennebec River. This difference in performance between the two assignment approaches is consistent with the prior stocking of the Kennebec River from Hudson River fish, as the underlying model behind Rubias has difficulty in discriminating populations with a large amount of admixture (Moran and Anderson 2019; LeBlanc et al. 2020). Given the similarity of the Kennebec River to the Hudson River and given that it likely does not contribute substantially to the mixed fishery, we recommend grouping the two rivers together in future mixed-stock analyses.

We ran mixture simulations to demonstrate the applicability of our SNP panel for mixed-stock analyses. Results showed highly accurate assignments to river of origin (92–100% for mixture sample sizes as low as 50 individuals), although there were outliers with low assignment rates. Given the presence of gene flow among the spawning rivers, especially within the Chesapeake Bay–Delaware River complex, it is unsurprising that there were individuals that could not be accurately assigned to a river by using the model in Rubias. This may also be the case if additional tributaries are added within the Chesapeake Bay. Nonetheless, the high median assignment accuracy for each reporting river indicates that our genetic panel would be useful for conducting mixed-stock analyses to identify the river of origin. River-level assignments can be improved by using AssignPOP as a follow-up analysis to identify individuals that cannot be accurately assigned with Rubias. Alternatively, if river-level assignments are not a priority, mixtures can be characterized with 100% accuracy by using reporting groups that combine rivers connected by gene flow (e.g., the Hudson–Kennebec River and the Chesapeake Bay–Delaware River complex).

Assignment accuracies did not vary with the number of individuals included in the mixture or with differing proportions of individuals in the mixture. Our genetic panel, therefore, is applicable to both small and large sampling efforts, providing an economically feasible tool for fishery managers. Given the highly mobile nature and differential recruitment success of Striped Bass (Goodyear and Christensen 1984; Ulanowicz and Polgar 1989; Rutherford and Houde 1995; Secor and Houde 1995; Secor 2000), it is likely that the composition of the coastal mixed fisheries changes temporally and spatially (Euclide et al. 2021). Our genetic panel is robust to this variation, as it showed consistently high accuracy of assignments regardless of which spawning river comprised the majority of the mixture. Reducing the number of individuals in the reference, however, reduced the accuracy of assignment for three rivers. This is not surprising and indicates that as many individuals as possible should be used to form a reference data set. At a minimum, 35 individuals should be used for locations with low genetic differentiation, while as few as 20 individuals could be used for locations that show strong signals of differentiation.

CONCLUSIONS

Striped Bass exhibit variability in their migratory behavior, including straying among rivers and skipped spawning (Kneebone et al. 2014; Callihan et al. 2015; Gahagan et al. 2015; Secor et al. 2020). Despite the highly vagile nature of Striped Bass, we found population differentiation at the level of individual rivers by using neutral and adaptive loci. Tailoring management actions to this fine spatial scale is important to protect against disproportional harvests of any particular population, especially the smaller contributors to mixed stocks (Cadrin and Secor 2009; Reiss et al. 2009; Kovach et al. 2010). Our study also highlights the importance of incorporating outlier loci and rare variants into population genetic analyses, as they can help to elucidate subtle patterns of differentiation. The population genetic structure is temporally stable, and the level of differentiation, while not large, is sufficient to assign individuals to river of origin. The panel of genetic markers developed in this study can be applied in future work via targeted sequence capture ("RADcap" approach; Hoffberg et al. 2016), thereby providing a high-resolution tool for accurate mixed-stock analyses and other management applications that will prove useful in light of the recent population declines of Striped Bass.

ACKNOWLEDGMENTS

We are grateful to New Hampshire Sea Grant for providing funding for this study (Project NA18OAR4170090). We also thank Isaac Wirgin for granting us access to his sample collection and for discussions about early results of the research. We are grateful for laboratory help from Rustam Bagirzadeh. We appreciate J. Clark, A. Scott, M. Bauer, L. Fenderson, L. Maxwell, and T. Kuras for feedback on an earlier version of the manuscript.

CONFLICT OF INTEREST

There are no conflicts of interest declared in this article.

ETHICS STATEMENT

This research followed the guidelines and was approved by the University of New Hampshire's Institutional Animal Care and Use Committee.

DATA AVAILABILITY STATEMENT

Genotype data generated in this paper are available upon request from the authors.

ORCID

Adrienne I. Kovach https://orcid.org/0000-0002-6791-0610

REFERENCES

Ackerman, M. W., Habicht, C. R., & Seeb, L. W. (2011). Single-nucleotide polymorphisms (SNPs) under diversifying selection provide increased accuracy and precision in mixed-stock analyses of Sockeye Salmon from the Copper River, Alaska. *Transactions of the American Fisheries Society*, *140*(3), 865–881. https://doi.org/10.1080/00028487.2011.588137

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697–709. https://doi.org/10.1038/nrg2844

Anderson, A. P., Denson, M. R., & Darden, T. L. (2014). Genetic structure of Striped Bass in the southeastern United States and effects from stock enhancement. *North American Journal of Fisheries Management*, *34*(3), 653–667. https://doi.org/10.1080/02755947.2014.902409

Anderson, E. C., Waples, R. S., & Kalinowski, S. T. (2008). An improved method for predicting the accuracy of genetic stock identification. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(7), 1475–1486. https://doi.org/10.1139/F08-049

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data* [Computer software]. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Andrews, S. N., Dadswell, M. J., Buhariwalla, C. F., Linnansaari, T., & Curry, R. A. (2019). Looking for Striped Bass in Atlantic Canada: The reconciliation of local, scientific, and historical knowledge. *Northeastern Naturalist*, *26*(1), 1–30. https://doi.org/10.1656/045.026.0105

ASMFC (Atlantic States Marine Fisheries Commission). (1981). *Interstate fisheries management plan for the Striped Bass* (Fisheries Management Report No. 1). ASMFC.

ASMFC (Atlantic States Marine Fisheries Commission). (2003). *Striped Bass technical committee advisory and summary reports on the status of the Atlantic Striped Bass*. ASMFC.

Avise, J., Arnold, J., Ball, M., Bermingham, E., Lamb, T., Neigel, J., Reeb, C., & Saunders, N. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, *19*, 489–522.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE*, *3*(10), Article e3376. https://doi.org/10.1371/journal.pone.0003376

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., & Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Molecular Ecology*, *24*(13), 3299–3315. https://doi.org/10.1111/mec.13245

Berggren, T. J., & Lieberman, J. T. (1978). Relative contribution of Hudson, Chesapeake, and Roanoke Striped Bass, (*Morone saxatilis*), stocks to the Atlantic coast fishery. *U.S. National Marine Fisheries Service Fishery Bulletin*, *76*(2), 335–345.

Bermingham, E., Rohwer, S., Freeman, S., & Wood, C. (1992). Vicariance biogeography in the Pleistocene and speciation in North American wood warblers: A test of Mengel's model. *Proceedings of the National Academy of Sciences of the USA*, *89*(14), 6624–6628. https://doi.org/10.1073/pnas.89.14.6624

Bielawski, J. P., & Pumo, D. E. (1997). Randomly amplified polymorphic DNA (RAPD) analysis of Atlantic Coast Striped Bass. *Heredity*, *78*, 32–40. https://doi.org/10.1038/hdy.1997.4

Boreman, J., & Austin, H. M. (1985). Production and harvest of anadromous Striped Bass stocks along the Atlantic coast. *Transactions of the American Fisheries Society*, *114*(1), 3–7. https://doi.org/10.1577/1548-8659(1985)114%3C3:PAHOAS%3E2.0.CO;2

Cadrin, S. X., & Secor, D. H. (2009). Accounting for spatial population structure in stock assessment: Past, present, and future. In R. J. Beamish & B. J. Rothschild (Eds.), *The future of fisheries science in North America* (pp. 405–426). Springer. https://doi.org/10.1007/978-1-4020-9210-7_22

Callihan, J. L., Godwin, C. H., & Buckel, J. A. (2014). Effect of demography on spatial distribution: Movement patterns of the Albemarle Sound–Roanoke River stock of Striped Bass (*Morone saxatilis*) in relation to their recovery. *U.S. National Marine Fisheries Service Fishery Bulletin*, *112*(2–3), 131–143. https://doi.org/10.7755/FB.112.2-3.3

Callihan, J. L., Harris, J. E., & Hightower, J. E. (2015). Coastal migration and homing of Roanoke River Striped Bass. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem*

Science, *7*(1), 301–315. https://doi.org/10.1080/19425120.2015.1057309

Campbell, E. O., Brunet, B. M. T., Dupuis, J. R., & Sperling, F. A. H. (2018). Would an RRS by any other name sound as RAD? *Methods in Ecology and Evolution*, *9*(9), 1920–1927. https://doi.org/10.1111/2041-210X.13038

Carvajal-Rodríguez, A. (2018). Myriads: *P*-value-based multiple testing correction. *Bioinformatics*, *34*(6), 1043–1045. https://doi.org/10.1093/bioinformatics/btx746

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Chen, K. Y., Euclide, P. T., Ludsin, S. A., Larson, W. A., Sovic, M. G., Gibbs, H. L., & Marschall, E. A. (2020). RAD-Seq refines previous estimates of genetic structure in Lake Erie Walleye. *Transactions of the American Fisheries Society*, *149*(2), 159–173. https://doi.org/10.1002/tafs.10215

Chen, K. Y., Marschall, E. A., Sovic, M. G., Fries, A. C., Gibbs, H. L., & Ludsin, S. A. (2018). AssignPOP: An R package for population assignment using genetic, non-genetic, or integrated data in a machine-learning framework. *Methods in Ecology and Evolution*, *9*(2), 439–446. https://doi.org/10.1111/2041-210X.12897

Dadswell, M. J., Spares, A. D., Mclean, M. F., Harris, P. J., & Rulifson, R. A. (2018). Long-term effect of a tidal, hydroelectric propeller turbine on the populations of three anadromous fish species. *Journal of Fish Biology*, *93*(2), 192–206. https://doi.org/10.1111/jfb.13755

Dahle, G., Johansen, T., Westgaard, J., Aglen, A., & Glover, K. A. (2018). Genetic management of mixed-stock fisheries 'real-time': The case of the largest remaining cod fishery operating in the Atlantic in 2007–2017. *Fisheries Research*, *205*, 77–85. https://doi.org/10.1016/j.fishres.2018.04.006

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., McVean, S. T., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Douglas, S. G., Bradford, R. G., & Chaput, G. (2003). *Assessment of Striped Bass* (Morone saxatilis) *in the maritime provinces in the context of species at risk*, 55 (Canadian Science Advisory Secretariat Research Document 2003/008). Fisheries and Oceans Canada.

Douglas, S. G., Chaput, G., Hayward, J., & Sheasgreen, J. (2009). Prespawning, spawning, and postspawning behavior of Striped Bass in the Miramichi River. *Transactions of the American Fisheries Society*, *138*(1), 121–134. https://doi.org/10.1577/T07-218.1

Dray, S., & Dufour, A. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*, Article 4. https://doi.org/10.18637/jss.v022.i04

Drinan, D. P., Gruenthal, K. M., Canino, M. L. F., Lowry, D., Fisher, M. C., & Hauser, L. (2018). Population assignment and local adaptation along an isolation-by-distance gradient in Pacific Cod (*Gadus macrocephalus*). *Evolutionary Applications*, *11*(8), 1448–1464. https://doi.org/10.1111/eva.12639

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Euclide, P. T., MacDougall, T., Robinson, J. M., Faust, M. D., Wilson, C. C., Chen, K., Marschall, E. A., Larson, W., & Ludsin, S. (2021). Mixed-stock analysis using Rapture genotyping to evaluate stock-specific exploitation of a Walleye population despite weak genetic structure. *Evolutionary Applications*, *14*(5), 1403–1420. https://doi.org/10.1111/eva.13209

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fabrizio, M. C. (1987). Contribution of Chesapeake Bay and Hudson River stocks of Striped Bass to Rhode Island coastal waters as estimated by isoelectric focusing of eye lens proteins. *Transactions of the American Fisheries Society*, *116(4)*, 588–593. https://doi.org/10.1577/1548-8659(1987)116%3C588:COCBAH%3E2.0.CO;2

Flagg, L. N., & Squires, T. S. (1994). *Restoration of Striped Bass in the state of Maine*. Maine Department of Marine Resources, Anadromous Fish Division.

Flannery, B. G., Beacham, T. D., Candy, J. R., Holder, R. R., Maschmann, G. F., Kretschmer, E. J., & Wenburg, J. K. (2010). Mixed-stock analysis of Yukon River Chum Salmon: Application and validation in a complex fishery. *North American Journal of Fisheries Management*, *30*(5), 1324–1338. https://doi.org/10.1577/M10-014.1

Gagnaire, P., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., & Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, *8*(8), 769–786. https://doi.org/10.1111/eva.12288

Gahagan, B. I., Fox, D. A., & Secor, D. H. (2015). Partial migration of Striped Bass: Revisiting the contingent hypothesis. *Marine Ecology Progress Series*, *525*, 185–197. https://doi.org/10.3354/meps11152

Gauthier, D. T., Audemard, C. A., Carlsson, J. E. L., Darden, T. L., Denson, M. R., Reece, K. S., & Carlsson, J. (2013). Genetic population structure of U.S. Atlantic coastal Striped Bass (*Morone saxatilis*). *Journal of Heredity*, *104*(4), 510–520. https://doi.org/10.1093/jhered/est031

Goodyear, C. P., & Christensen, S. W. (1984). On the ability to detect the influence of spawning stock on recruitment. *North American Journal of Fisheries Management*, *4*(2), 186–193. https://doi.org/10.1577/1548-8659(1984)4<186:OTATDT>2.0.CO;2

Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., Martino, J. A., Pierson, T., Rogers, S. M., Wilson, J. Y., & Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, *15*(6), 1304–1315. https://doi.org/10.1111/1755-0998.12404

Hess, J. E., Matala, A. P., & Narum, S. R. (2011). Comparison of SNPs and microsatellites for fine-scale application of genetic stock identification of Chinook Salmon in the Columbia River basin. *Molecular Ecology Resources*, *11*(s1), 137–149. https://doi.org/10.1111/j.1755-0998.2010.02958.x

Hilborn, R., Quinn, T. P., Schindler, D. E., & Rogers, D. E. (2003). Biocomplexity and fisheries sustainability. *Proceedings of the National Academy of Sciences of the USA*, *100*(11), 6564–6568. https://doi.org/10.1073/pnas.1037274100

Hoffberg, S. L., Kieran, T. J., Catchen, J. M., Devault, A., Faircloth, B. C., Mauricio, R., & Glenn, T. C. (2016). RADcap: Sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, *16*(5), 1264–1278. https://doi.org/10.1111/1755-0998.12566

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*(5), 1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Jenkins, T. L., Ellis, C. D., Triantafyllidis, A., & Stevens, J. R. (2019). Single nucleotide polymorphisms reveal a genetic cline across the north-east Atlantic and enable powerful population assignment in the European lobster. *Evolutionary Applications*, *12*(10), 1881–1899. https://doi.org/10.1111/eva.12849

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jorde, P. E., Synnes, A., Espeland, S. H., Sodeland, M., & Knutsen, H. (2018). Can we rely on selected genetic markers for population identification? Evidence from coastal Atlantic Cod. *Ecology and Evolution*, *8*(24), 12547–12558. https://doi.org/10.1002/ece3.4648

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, Article e281. https://doi.org/10.7717/peerj.281

Kenter, L. W., Kovach, A. I., Woods, L. C., Reading, B. J., & Berlinsky, D. L. (2018). Strain evaluation of Striped Bass (*Morone saxatilis*) cultured at different salinities. *Aquaculture*, *492*, 215–225. https://doi.org/10.1016/j.aquaculture.2018.04.017

Kerr, L. A., Hintzen, N. T., Cadrin, S. X., Clausen, L. W., Dickey-Collas, M., Goethel, D. R., Hatfield, E. M. C., Kritzer, J. P., & Nash, R. D. M. (2017). Lessons learned from practical approaches to reconcile mismatches between biological population structure and stock units of marine fish. *ICES Journal of Marine Science*, *74*(6), 1708–1722. https://doi.org/10.1093/icesjms/fsw188

Kneebone, J., Hoffman, W. S., Dean, M. J., Fox, D. A., & Armstrong, M. P. (2014). Movement patterns and stock composition of adult Striped Bass tagged in Massachusetts coastal waters. *Transactions of the American Fisheries Society*, *143*(5), 1115–1129. https://doi.org/10.1080/00028487.2014.889752

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Kovach, A. I., Breton, T. S., Berlinsky, D. L., Maceda, L., & Wirgin, I. (2010). Fine-scale spatial and temporal genetic structure of Atlantic Cod off the Atlantic coast of the USA. *Marine Ecology Progress Series*, *410*, 177–195. https://doi.org/10.3354/meps08612

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with B bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Leblanc, N. M., Andrews, S. N., Avery, T. S., Puncher, G. N., Gahagan, B. I., Whiteley, A. R., Curry, R., & Pavey, S. A. (2018). Evidence of a genetically distinct population of Striped Bass within the Saint John River, New Brunswick, Canada. *North American Journal of Fisheries Management*, *38*(6), 1339–1349. https://doi.org/10.1002/nafm.10242

LeBlanc, N. M., Gahagan, B. I., Andrews, S. N., Avery, T. S., Puncher, G. N., Reading, B. J., Buhariwalla, C. F., Curry, R., Whiteley, A. R., & Pavey, S. A. (2020). Genomic population structure of Striped Bass (*Morone saxatilis*) from the Gulf of St. Lawrence to Cape Fear River. *Evolutionary Applications*, *13*(6), 1468–1486. https://doi.org/10.1111/eva.12990

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Luu, K., Bazin, E., & Blum, M. G. B. (2017). Pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, *17*(1), 67–77. https://doi.org/10.1111/1755-0998.12592

Martinez, S., Willoughby, J. R., & Christie, M. R. (2018). Genetic diversity in fishes is influenced by habitat type and life-history variation. *Ecology and Evolution*, *8*(23), 12022–12031. https://doi.org/10.1002/ece3.4661

Martinsohn, J. T., & Ogden, R. (2009). FishPopTrace—Developing SNP-based population genetic assignment methods to investigate illegal fishing. *Forensic Science International: Genetics Supplement Series*, *2*(1), 294–296. https://doi.org/10.1016/j.fsigss.2009.08.108

McLean, J., & Taylor, E. (2001). Resolution of population structure in a species with high gene flow: Microsatellite variation in the eulachon (Osmeridae: *Thaleichthys pacificus*). *Marine Biology*, *139*(3), 411–420. https://doi.org/10.1007/s002270100483

Meirmans, P. G. (2020). Genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources*, *20*(4), 1126–1131. https://doi.org/10.1111/1755-0998.13145

Meirmans, P. G., & Liu, S. (2018). Analysis of molecular variance (AMOVA) for autopolyploids. *Frontiers in Ecology and Evolution*, *6*, Article 66. https://doi.org/10.3389/fevo.2018.00066

Moran, B. M., & Anderson, E. C. (2019). Bayesian inference from the conditional genetic stock identification model. *Canadian Journal of Fisheries and Aquatic Sciences*, *76*(4), 551–560. https://doi.org/10.1139/cjfas-2018-0016

National Marine Fisheries Service. (2020). *Fisheries of the United States, 2020* (National Oceanic and Atmospheric Administration Current Fishery Statistics No. 2020). National Marine Fisheries Service. https://www.fisheries.noaa.gov/resource/document/fisheries-united-states-2020

NEFSC (Northeast Fisheries Science Center). (2019). *66th Northeast regional Stock Assessment Workshop (66th SAW) assessment report* (NEFSC Reference Document 19-08). NEFSC. https://doi.org/10.25923/nhqe-jd35

Nielsen, E. E., Cariani, A., Aoidh, E. M., Maes, G. E., Milano, I., Ogden, R., Taylor, M., Hemmer-Hansen, J., Babbucci, M., Bargelloni, L., Bekkevold, D., Diopere, E., Grenfell, L., Helyar, S., Limborg, M. T., Martinsohn, J. T., McEwing, R., Panitz, F., Patarnello, T., … Carvalho, G. R. (2012). Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications*, *3*, Article 851. https://doi.org/10.1038/ncomms1845

Nielsen, E. E., Hemmer-Hansen, J., Poulsen, N. A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G., Ogden, R., & Carvalho, G. R. (2009). Genomic signatures of local directional selection in a high gene flow marine organism: The Atlantic Cod (*Gadus morhua*). *BMC Evolutionary Biology*, *9*(1), Article 276. https://doi.org/10.1186/1471-2148-9-276

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you're looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*(16), 2193–3206. https://doi.org/10.1111/mec.14792

Palsboll, P., Berube, M., & Allendorf, F. (2007). Identification of management units using population genetic data. *Trends in Ecology & Evolution*, *22*(1), 11–16. https://doi.org/10.1016/j.tree.2006.09.003

Paradis, E. (2010). Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reiss, H., Hoarau, G., Dickey-Collas, M., & Wolff, W. J. (2009). Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries*, *10*(4), 361–395. https://doi.org/10.1111/j.1467-2979.2008.00324.x

Robinson, M., Courtenay, S., Benfey, T., Maceda, L., & Wirgin, I. (2004). Origin and movements of young-of-the-year Striped Bass in the southern Gulf of St. Lawrence, New Brunswick. *Transactions of the American Fisheries Society*, *133*(2), 412–426. https://doi.org/10.1577/01-073

Robitaille, J., Bérubé, M., Gosselin, A., Baril, M., Beauchamp, J., Boucher, J., Dionne, S., Legault, M., Mailhot, Y., Ouellet, B., Sirois, P., Tremblay, S., Trencia, G., Verreault, G., & Villeneuve, D. (2011). *Recovery strategy for the Striped Bass* (Morone saxatilis)*, St. Lawrence Estuary population, Canada*. Fisheries and Oceans Canada. http://epe.lac-bac.gc.ca/100/201/301/weekly_checklist/2012/internet/w12-03-U-E.html/collections/collection_2012/mpo-dfo/En3-4-105-2011-eng.pdf

Rothermel, E. R., Balazik, M. T., Best, J. E., Breece, M. W., Fox, D. A., Gahagan, B. I., Haulsee, D. E., Higgs, A. L., O'Brien, M. H. P., Oliver, M. J., Park, I. A., & Secor, D. H. (2020). Comparative migration ecology of Striped Bass and Atlantic Sturgeon in the US Southern mid-Atlantic bight flyway. *PLOS ONE*, *15*(6), Article e0234442. https://doi.org/10.1371/journal.pone.0234442

Rulifson, R. A., & Dadswell, M. J. (1995). Life history and population characteristics of Striped Bass in Atlantic Canada. *Transactions of the American Fisheries Society*, *124*(4), 477–507. https://doi.org/10.1577/1548-8659(1995)124%3C0477:LHAPCO%3E2.3.CO;2

Rulifson, R. A., McKenna, S. A., & Dadswell, M. J. (2008). Intertidal habitat use, population characteristics, movement, and exploitation of Striped Bass in the inner Bay of Fundy, Canada. *Transactions of the American Fisheries Society*, *137*(1), 23–32. https://doi.org/10.1577/T06-174.1

Russello, M. A., Kirk, S. L., Frazer, K. K., & Askey, P. J. (2012). Detection of outlier loci and their utility for fisheries management: Outlier loci for fisheries management. *Evolutionary Applications*, *5*(1), 39–52. https://doi.org/10.1111/j.1752-4571.2011.00206.x

Rutherford, E. S., & Houde, E. D. (1995). The influence of temperature on cohort-specific growth, survival, and recruitment of Striped Bass (*Morone saxatilis*) larvae in Chesapeake Bay. *U.S. National Marine Fisheries Service Fishery Bulletin*, *93*, 315–332.

Schneider, C. J., Cunningham, M., & Moritz, C. (1998). Comparative phylogeography and the history of endemic vertebrates in the wet tropics rainforests of Australia. *Molecular Ecology*, *7*(4), 487–498. https://doi.org/10.1046/j.1365-294x.1998.00334.x

Secor, D. H. (2000). Spawning in the nick of time? Effect of adult demographics on spawning behaviour and recruitment in Chesapeake Bay Striped Bass. *ICES Journal of Marine Science*, *57*(2), 403–411. https://doi.org/10.1006/jmsc.1999.0520

Secor, D. H., & Houde, E. D. (1995). Temperature effects on the timing of Striped Bass egg production, larval viability, and recruitment potential in the Patuxent River (Chesapeake Bay). *Estuaries*, *18*(3), 527–544. https://doi.org/10.2307/1352370

Secor, D. H., O'Brien, M. H., Gahagan, B. I., Watterson, J. C., & Fox, D. A. (2020). Differential migration in Chesapeake Bay Striped Bass. *PLOS ONE*, *15*(5), Article e0233103. https://doi.org/10.1371/journal.pone.0233103

Secor, D. H., & Piccoli, P. M. (2007). Oceanic migration rates of upper Chesapeake Bay Striped Bass (*Morone saxatilis*), determined by otolith microchemical analysis. *U.S. National Marine Fisheries Service Fishery Bulletin*, *105*(1), 62–73.

Stapley, J., Reger, J., Feulner, P. G. D., Smadja, C., Galindo, J., Ekblom, R. T., Bennison, C., Ball, A. D., Beckerman, A. P., & Slate, J. (2010). Adaptation genomics: The next generation. *Trends in Ecology & Evolution*, *25*(12), 705–712. https://doi.org/10.1016/j.tree.2010.09.002

Storey, J. D., Bass, A. J., Dabney, A., & Robinson, D. (2022). *qvalue: Q-value estimation for false discovery rate control*. R package version 2.30.0. http://github.com/jdstorey/qvalue

Sutherland, B. J. G., Candy, J., Mohns, K., Cornies, O., Jonsen, K., Le, K., Gustafson, R. G., Nichols, K. M., & Beacham, T. D. (2021). Population structure of Eulachon (*Thaleichthys pacificus*) from Northern California to Alaska using single nucleotide polymorphisms from direct amplicon sequencing. *Canadian Journal of Fisheries and Aquatic Sciences*, *78*(1), 78–89. https://doi.org/10.1139/cjfas-2020-0200

Ulanowicz, R. E., & Polgar, T. T. (1989). Influences of anadromous spawning behaviour and optimal environmental conditions upon Striped Bass (*Morone saxatilis*) year-class strength. *Canadian Journal of Fisheries and Aquatic Sciences*, *37(2)*, 143–154. https://doi.org/10.1139/f80-019

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer. https://www.stats.ox.ac.uk/pub/MASS4/

Vendrami, D. L. J., Telesca, L., Weigand, H., Weiss, M., Fawcett, K., Lehman, K., Clark, M. S., Leese, F., McMinn, C., Moore, H., & Hoffman, J. I. (2017). RAD sequencing resolves fine-scale population structure in a benthic invertebrate: Implications for understanding phenotypic plasticity. *Royal Society Open Science*, *4*(2), Article 160548. https://doi.org/10.1098/rsos.160548

Waldman, J., Maceda, L., & Wirgin, I. (2012). Mixed-stock analysis of wintertime aggregations of Striped Bass along the mid-Atlantic coast. *Journal of Applied Ichthyology*, *28*(1), 1–6. https://doi.org/10.1111/j.1439-0426.2011.01888.x

Waldman, J. R., Dunning, D. J., Ross, Q. E., & Mattson, M. T. (1990). Range dynamics of Hudson River Striped Bass along the Atlantic coast. *Transactions of the American Fisheries Society*, *119*(5), 910–919. https://doi.org/10.1577/1548-8659(1990)119%3C0910:RDOHRS%3E2.3.CO;2

Waldman, J. R., & Wirgin, I. I. (1995). Origin of the present Delaware River Striped Bass population as shown by analysis of mitochondrial DNA. *Transactions of the American*

Fisheries Society, *123*(1), 15–21. https://doi.org/10.1577/1548-8659(1994)123%3C0015:OOTPDR%3E2.3.CO;2

Ward, R. D. (2006). The importance of identifying spatial population structure in restocking and stock enhancement programmes. *Fisheries Research*, *80*(1), 9–18. https://doi.org/10.1016/j.fishres.2006.03.009

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of F_{ST}. *American Naturalist*, *186*(S1), S24–S36. https://doi.org/10.1086/682949

Wirgin, I., Maceda, L., Tozer, M., Stabile, J., & Waldman, J. (2020). Atlantic coastwide population structure of Striped Bass (*Morone saxatilis*) using microsatellite DNA analysis. *Fisheries Research*, *226*, Article 105506. https://doi.org/10.1016/j.fishres.2020.105506

Wirgin, I., Maceda, L., Waldman, J. R., & Crittenden, R. N. (1993). Use of mitochondrial DNA polymorphisms to estimate relative contributions of the Hudson River and Chesapeake Bay Striped Bass stocks to mixed fishery on the Atlantic coast. *Transactions of the American Fisheries Society*, *122*(5), 669–684. https://doi.org/10.1577/1548-8659(1993)122%3C0669:UOMDPT%3E2.3.CO;2

Wirgin, I. I., Silverstein, P., & Grossfield, J. (1990). Restriction endonuclease analysis of Striped Bass mitochondrial DNA: The Atlantic coastal migratory stock. In N. C. Parker, A. E. Giorgi, R. C. Heidinger, D. B. Jester Jr, E. D. Prince, & G. A. Winans (Eds.), *Fish-marking techniques* (Symposium 7, pp. 475–491). American Fisheries Society.

Wirgin, I. I., Waldman, J. R., Maceda, L., Stabile, J., & Vecchio, V. J. (1997). Mixed-stock analysis of Atlantic coast Striped Bass (*Morone saxatilis*) using nuclear DNA and mitochondrial DNA markers. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*(12), 2814–2826. https://doi.org/10.1139/f97-195

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.