

Disentangling complex genomic signals to understand population structure of an exploited, estuarine-dependent flatfish

Shannon J. O'Leary¹  | Christopher M. Hollenbeck² | Robert R. Vega³ | David S. Portnoy²

¹Department of Biology, Saint Anselm College, Manchester, New Hampshire, USA

²Marine Genomics Laboratory, Department of Life Sciences, Texas A&M University Corpus Christi, Corpus Christi, Texas, USA

³CCA Marine Development Center, Texas Parks and Wildlife Department, Corpus Christi, Texas, USA

Correspondence

Shannon J. O'Leary, Department of Biology, Saint Anselm College, 100 Saint Anselm Drive, Manchester, NH 03102, USA.
Email: shannon.j.oleary@gmail.com

Funding information

State Wildlife Grant Program, Grant/Award Number: CFDA# 15.634

Abstract

Interpreting contemporary patterns of population structure requires an understanding of the interactions among microevolutionary forces and past demographic events. Here, 4,122 SNP-containing loci were used to assess structure in southern flounder (*Paralichthys lethostigma*) sampled across its range in the US Atlantic Ocean (Atlantic) and Gulf of Mexico (Gulf) and relationships among components of genomic variation and spatial and environmental variables were assessed across estuarine population samples in the Gulf. While hierarchical AMOVA revealed significant heterogeneity within and between the Atlantic and Gulf, pairwise comparisons between samples within ocean basins demonstrated that all significant heterogeneity occurred within the Gulf. The distribution of Tajima's *D* estimated at a genome-wide scale differed significantly from equilibrium in all estuaries, with more negative values occurring in the Gulf. Components of genomic variation were significantly associated with environmental variables describing individual estuaries, and environment explained a larger component of variation than spatial proximity. Overall, results suggest that there is genetic spatial autocorrelation caused by shared larval sources for proximal nurseries (migration/drift), but that it is modified by environmentally driven differentiation (selection). This leads to conflicting signals in different parts of the genome and creates patterns of divergence that do not correspond to paradigms of strong local directional selection.

KEYWORDS

genetic-environmental association, local adaptation, population genomics

1 | INTRODUCTION

Patterns of contemporary genetic variation within and among populations result from interactions among microevolutionary forces (i.e., genetic drift, migration, selection, and mutation) against the

backdrop of historical demographic changes, including population expansion and bottlenecks (Luikart et al., 2003). Because these forces result in allele frequency changes at specific loci (e.g., selection) and across loci in the genome as a whole (e.g., drift), studies employing dense genome-wide sampling of genetic markers have

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

become the gold standard for understanding and interpreting current patterns of population structure in exploited marine species (Bernatchez et al., 2017). Central to this endeavor is a model, supported by both population genetic theory and empirical studies, that posits strong geographically localized directional selection which will create elevated divergence with respect to the rest of the genome and that these areas of increased divergence (outlier loci) should be confined to a small proportion of the genome (Lewontin & Krakauer, 1973; Nielsen, 2001).

While this paradigm is important, other factors may confound interpretation, including selection pressures changing with ontogeny, polygenic traits, violations of equilibrium assumptions, and complex interactions of other microevolutionary forces (Forester et al., 2018; Hoban et al., 2016; Lotterhos & Whitlock, 2014). Consideration of such factors may be particularly important for marine bony fishes, which frequently feature complex life histories in which the larval phase differs greatly in size, behavior, and other aspects of basic biology from the juvenile and adult phases. While juveniles are more similar to adults than larvae, they usually occupy different trophic levels and have different habitat requirements than adults. Habitat changes may occur several times during development, suggesting ontogenetic shifts in selective pressures are likely (Dahlgren & Eggleston, 2000; Yang et al., 2018). Additionally, many marine species have large population sizes and high fecundity (Hedgecock & Pudovkin, 2011). This results in a decrease of the relative strength of genetic drift while increasing the opportunity for adaptive variants to arise via mutation (Cormack et al., 1990). Further, marine species are frequently geographically widespread across heterogeneous environments, leading to opportunities for localized selection to create adaptive genetic variation underlying phenotypic traits including life history, morphology, behavior, and physiology (Bernatchez, 2016). Frequently, these are polygenic traits, though some complex traits may also be controlled by a single locus or chromosomal region (Prince et al., 2017). High degrees of connectivity due to larval dispersal and adult migration are typical of marine species and while theoretically this could weaken the effects of selection (Felsenstein, 1976), studies have shown that marine species often display local adaptation despite high gene flow (Clarke et al., 2010; Hoey & Pinsky, 2018). Taken together, the combination of these factors could lead to a scenario where selection acts upon different genes and regions of the genome with a wide range of effect sizes (Gagnaire & Gaggiotti, 2016) that are not always detectable using outlier detection (genome scan) methods (Bernatchez, 2016; Gagnaire & Gaggiotti, 2016). Finally, for many marine species, historical changes in climate and sea level have led to fluctuations in population size and connectivity (Marko & Hart, 2012; Portnoy et al., 2014), which may lead to remnant historical demographic signals still present in the genome.

Southern flounder, *Paralichthys lethostigma*, inhabit estuarine and nearshore environments along the U.S. Atlantic coast (Atlantic from the Carolinas to Florida and across the northern Gulf of Mexico (Gulf) to near Veracruz in Mexican waters, with a break in distribution along the southern Florida peninsula. They support substantial commercial and recreational fisheries, accounting for a

multi-million-dollar fishery in U.S. waters of the Atlantic and northern Gulf (Flounder Technical Task Force, 2015). Long-term declines in abundance of young-of-the-year (YOY), juvenile, and adult southern flounder are well documented in the Gulf and have been attributed to fishing mortality (Flounder Technical Task Force, 2015; Froeschke et al., 2011). Recently, steep declines in the number of YOY in the western Gulf have prompted an interest in a stock-enhancement program where hatchery-reared YOY, spawned from captive wild-caught adults, are used to augment recruitment (Kaiser et al., 2012; Miller et al., 2010). Both the high market value of southern flounder and its suitability as a target for commercial and stock augmentation (Daniels et al., 1996; Watanabe et al., 2006) warrant a robust assessment of population structure.

Southern flounder have a high dispersal potential resulting from larval dispersal, ontogenetic shifts in habitat use, and adult movement. Adults reside in bays and estuaries for much of the year but migrate to spawning grounds located in multiple offshore locations in the fall. This provides opportunity for gene flow as spawning grounds do not necessarily correspond to individual estuaries of origin for adults (Craig et al., 2015). In addition, individual estuaries may receive larval subsidies from more than one spawning ground as buoyant eggs are transported to nursery habitats where larvae settle in shallow estuarine and freshwater environments in late winter and early spring (Nims & Walther, 2014). Larval supply to specific estuaries is affected by multiple processes, including currents, tides, and hydrography of individual estuaries. In addition, long larval duration (30–60 days) results in the potential for long-range dispersal (Bailey et al., 2005). Southern flounder YOY exhibit shifts in habitat preference as they grow, and different size/age classes are found in different microhabitats within bays and estuaries, characterized by different levels of salinity and differences in substrate and vegetation (Furey & Rooker, 2013; Glass et al., 2008; Nañez-James et al., 2009). Juveniles remain in the same estuary until maturity at approximately two years of age, and tagging studies indicate limited adult movement during seasonal estuarine residence (Craig et al., 2015; Monaghan, 1992). Analyses of otolith microchemistry suggest a lack of nursery-site fidelity among adults (Wang et al., 2018), and little is known about habitat use during and after offshore spawning. Additionally, there is evidence that some adults may remain permanently offshore (Watterson & Alexander, 2004).

Previous studies of stock structure based on allozymes (Blandon et al., 2001), mitochondrial DNA (Anderson et al., 2012), microsatellites (Wang et al., 2015), and otolith morphometrics (Midway et al., 2014) have identified differences between southern flounder in the Atlantic and Gulf but failed to demonstrate differences within each region. By contrast, regional and estuarine-specific differences in life-history traits such as growth rates and age/size-at-maturity have been documented, suggesting the potential for genetic differences at sufficiently small spatial-scales that microsatellites and mtDNA cannot resolve (Corey et al., 2017; Fischer & Thompson, 2004). Additionally, heterogeneity in habitat within and among estuaries inhabited by southern flounder may impact growth and survival at multiple life stages (Corey et al., 2017; Midway

et al., 2015). This heterogeneity includes differences in habitat availability, levels of freshwater input, substrate types, hydrography, temperature, salinity, nutrient loading, and suspended sediment (NEAA, 2018). Variance in environmental conditions across estuaries could create differences in local selection pressures thus leading to changes in components of genetic variation that could persist despite gene flow or even contribute to the disruption of gene flow (Tigano & Friesen, 2016). Finally, historic events are important to shaping contemporary genomic variation. For example, population expansion after glacial periods have been demonstrated for multiple marine taxa in the Atlantic and Gulf (Marko & Hart, 2012; Portnoy et al., 2014) and are likely important for southern flounder as well.

To understand the relative roles that environment, geography, and demographic history have played in shaping contemporary population structure of southern flounder, reduced representation DNA sequencing was used to examine genomic variation of the southern flounder genome from individuals sampled throughout their range. By using several thousand loci randomly distributed throughout the genome, finer-scale patterns of population structure can be resolved as compared with previous data sets consisting of tens of loci (mtDNA/microsatellite studies). Additionally, these data sets consist of loci in the coding and noncoding parts of the genome, enabling an assessment of population structure using both neutral and presumably adaptive genetic variation and screening for associations among components of genomic variation and environmental variables. Typically, multiple single-nucleotide polymorphisms (SNPs) are detected on single DNA fragments generated using next-generation sequencing. To reduce effects of linkage, these SNPs need to be

thinned. Instead, a haplotype-based approach was implemented, generating a data set consisting of multiallelic loci, and thus combining both the power of a larger number of loci and multiple alleles per locus (Baetscher et al., 2018; Willis et al., 2017). Additionally, this approach allowed for the implementation of DNA sequence-based analyses to test for conformance to equilibrium assumptions at individual locus and genome-wide scales indicative of selection/local adaptation versus population expansion, respectively.

2 | METHODS

2.1 | Sampling design

Tissues (fin clips) were taken from YOY, juveniles, and adults sampled from six estuaries in the Atlantic and nine estuaries in the Gulf (Figure 1, Table 1) during regularly occurring surveys by the Florida Fish and Wildlife Conservation Commission, Texas Parks and Wildlife Department, and South Carolina Department of Natural Resources from 2013 to 2016. Because individuals remain in their natal estuaries for approximately two years, attempts were made to obtain exclusively YOY (defined as individuals caught at <25 cm) from each year to determine estuary of origin. Drastic declines in the number of YOY caught during regular surveys did not allow for sufficient sample sizes. Instead, sample sizes were supplemented with juveniles (25–40 cm) and adults (>40 cm) caught in or near estuaries. Small sample sizes for each cohort did not allow for a formal test of cohort effects

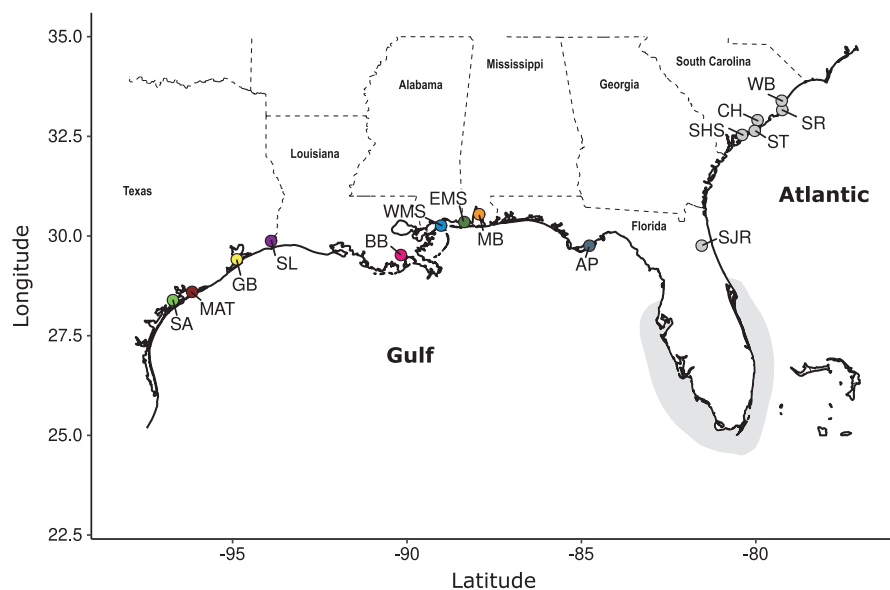


FIGURE 1 Sample distribution of southern flounder young-of-the-year, juveniles, and adults sampled in estuaries throughout the Gulf of Mexico and Western Atlantic. Sample locations from West to East and South to North are San Antonio Bay (SA, $N = 23$), Matagorda Bay (MAT, $N = 3$), Galveston Bay (GB, $N = 2$), Sabine Lake (SL, $N = 24$), Barataria Bay (BB, $N = 39$), West Mississippi Sound (WMS, $N = 34$), East Mississippi Sound (EMS, $N = 5$), Mobile Bay (MB, $N = 62$), Apalachicola Bay (AP, $N = 44$), St. John's River (SJR, $N = 20$), St. Helena Sound (SHS, $N = 4$), Stono-North Rivers ($N = 4$), Charleston Harbor (CH, $N = 18$), Santee Rivers (SR, $N = 4$), and Winyah Bay (WB, $N = 18$). Sample locations in the Gulf are colored to match Figures 4 and 5. The approximate extent of the break in distribution off the coast of Florida is indicated in grey

Ocean basin	Estuary	Adults	Juveniles	YOY
Gulf	Apalachicola Bay	1	3	40
	Barataria Bay	7	32	n/a
	Charleston Harbor	7	9	2
	East Mississippi Sound	3	2	n/a
	Galveston Bay	n/a	2	n/a
	Matagorda Bay	n/a	n/a	3
	Mobile Bay	48	14	n/a
	Sabine Lake	2	19	3
	San Antonio Bay	3	10	10
Atlantic	Santee Rivers	2	2	n/a
	St Helena Sound	2	2	n/a
	St John's River	n/a	n/a	20
	Stono-North Edisto Rivers	1	n/a	3
	West Mississippi Sound	9	23	2
	Winyah Bay	7	8	3

TABLE 1 Sample size per estuary and age class. Young of the year (YOY) defined as fish caught at 25 cm or less, juveniles as 25–40 cm and adults as >40 cm

and assignment of adults to estuaries, though exploratory analysis restricting the data set to only YOY caught in the same year yielded the same overall patterns as analysis with mixed cohorts, indicating that cohort effects are negligible. In addition, mixed age classes should help to minimize temporal effects caused by recruitment variation. Therefore, samples were pooled across life stages during formal data analysis.

2.2 | Genotyping

DNA was extracted using Mag-Bind Blood and Tissue DNA kits (Omega Bio-Tek). Double digest restriction site-associated DNA (ddRAD) libraries were constructed using a modified protocol (Portnoy et al., 2015) and sequenced on four separate lanes of an Illumina HiSeq 2500. Raw sequences were demultiplexed using *process_radtags* (Catchen et al., 2011). Quality trimming, read mapping, and SNP calling were performed using the *dDocent* pipeline (Puritz et al., 2014) and a reduced-representation reference genome (approximately 2%–5% of the genome) previously produced for southern flounder (O'Leary et al., 2018). Raw SNPs were rigorously filtered for quality following recommended practices (O'Leary et al., 2018). Quite commonly, more than one SNP was identified on a single DNA fragment, and rather than thin SNPs to retain only one biallelic SNP per contig, *rad_haplotyper* (Willis et al., 2017) was used to collapse SNPs on the same contig into haplotypes, producing a data set consisting of 4,122 multiallelic loci (2–80 alleles per locus). Data analysis was primarily allele frequency-based, apart from the test of neutrality using Tajima's *D*, which requires DNA sequences to test for mutation-drift equilibrium. Detailed processing steps, reproducible code, and threshold values used are available from https://github.com/sjoleary/SFL_PopGen.

2.3 | F_{ST} -outlier analysis

Presence of F_{ST} -outlier loci was assessed using two methods: the *FDIST*-method implemented in *ARLEQUIN* (Excoffier & Lischer, 2010), and the Bayesian approach used in *BAYESCAN* (Foll & Gaggiotti, 2008). For both methods, outlier loci with significantly higher F_{ST} values than expected under a neutral model are flagged as loci putatively under directional selection. Given low background F_{ST} values typically found in marine fishes (Waples, 1998), an assessment for loci putatively under balancing selection (F_{ST} significantly lower than expected) was not conducted. Analysis in *ARLEQUIN* was based on 20,000 coalescent simulations, using a strict island model. To account for multiple testing, *p*-values were corrected according to Benjamini and Hochberg (1995). *BAYESCAN* runs consisted of 25 pilot runs of 5,000 iterations, followed by a total of 550,000 iterations (burn-in of 50,000 iterations, 10,000 samples with a thinning interval of 50) a *q*-value of .05 was used as the threshold for outlier detection. For both methods, F_{ST} -outlier analysis was run using all individuals grouped by estuary and then using individuals grouped by estuary separately for each ocean basin. The distribution of loci flagged during F_{ST} -outlier analysis across linkage groups (chromosomes) was assessed using a previously established linkage map (O'Leary et al., 2018).

2.4 | Assessment of population structure and genetic diversity

Loci were subdivided into two datasets: outlier (consensus loci identified by both outlier detection methods) and neutral loci (all remaining loci). Hierarchical, locus-by-locus analysis of molecular variance (*AMOVA*), as implemented in *ARLEQUIN*, was used to test for homogeneity of genetic variation both between basins and among estuaries within basins. Homogeneity within each basin was explored for

neutral loci further using a single-level AMOVA. For AMOVA, the significance of each component of variation was assessed by permuting individuals between groups 10,000 times. Pairwise estimates of F_{ST} were also generated in ARLEQUIN, as a post hoc test for homogeneity between estuaries. Significance was determined as above and correction for multiple comparisons applied following Benjamini and Hochberg. Only samples from estuaries with ≥ 18 individuals were used in the analysis. Although 20–30 individuals has been suggested as the minimum sample size for reliable allele frequency estimates, choosing 18 samples allowed for balance of retaining as many groups as possible while providing sufficient power (Luikart & Cornuet, 1998; Morin et al., 2009).

Genomic diversity of each estuary was determined as Nei's gene diversity (Nei, 1973), rarefied allele counts, and evenness. The last is a measure of the distribution of allele frequencies and was estimated as the ratio of the Stoddart & Taylor index (diversity weighted for more abundant alleles) and the Shannon-Wiener index (diversity weighted for rarer alleles), as implemented in *poppr* (Kamvar et al., 2014). For each measure of diversity, a Friedman's rank sum test was used to test homogeneity among estuaries. A Wilcoxon signed-rank test was used post hoc to test for pairwise differences between estuaries; *p*-values were corrected for multiple comparisons according to Benjamini and Hochberg. The number of fixed loci was documented for each ocean basin and estuary and the global rarefied allele counts for loci fixed in a given group of individuals compared.

2.5 | Tests of neutrality at the genome and locus level

The observed genome-wide distribution of Tajima's *D* was calculated for each estuary and compared with a null-distribution of an equivalent set of loci in drift-mutation equilibrium. While significant deviations from equilibrium at individual loci are indicative of selection, a genome-wide deviation indicates that demographic events (population expansion/bottlenecks) have occurred. This analysis made use of the fact that DNA sequence information can be retrieved from microhaplotypes (SNP-containing loci). The population-scaled mutation rate Θ was estimated for each locus as the average number of pairwise differences per site between all pairs of microhaplotypes (Θ_T , nucleotide diversity; Nei, 1987), and as the number of segregating sites across all microhaplotypes (Watterson's estimator Θ_W , Watterson, 1975). In an equilibrium population of stable size and without selection, Θ_T will approximate Θ_W , but when a population has undergone recent expansion or there is strong directional selection, Θ_W will exceed Θ_T because the latter is sensitive to allele frequencies and the former is not. Tajima (1989) formalized this observation in a test statistic, Tajima's *D*, that can be used to test for conformance to drift-mutation equilibrium assumptions.

Tajima's *D* was calculated for each locus with individuals grouped by estuary, creating a set of estuary-specific observed distributions. Then, 1,000 genome-wide, null distributions of Tajima's *D* were

simulated for each estuary, using a coalescent model, executed in MS (Hudson, 2002). To do this, a set of neutral loci consisting of the same number of loci with the same distribution of segregating sites as in the observed data (grouped by estuary) was generated. For example, if 100 loci in the empirical data set have four segregating sites, the simulated data set also contains 100 loci with four segregating sites. The difference between observed and simulated distributions was then assessed for each estuary by comparing mean and median values of the empirical distribution with those of the simulated distributions. Significance was assessed by determining the proportion of times that the observed value was smaller or larger than simulated values.

Further, locus-specific effects were assessed by testing for significant deviation from neutrality as implemented in *pegas* (Paradis, 2010); *p*-values were corrected by estuary for each locus according to Benjamini and Hochberg (1995) to account for multiple comparisons. The number of significant ($p < .05$) loci, positive and negative, was then summarized by estuary. The distribution of loci with significant Tajima's *D* values across linkage groups was assessed using only those loci that were previously incorporated into a linkage map (O'Leary et al., 2018).

Finally, to better understand what was driving patterns in neutrality tests, Θ was calculated based on the number of segregating sites as Θ_W (Watterson, 1975) and based on pairwise differences among haplotypes as Θ_T (nucleotide diversity; Nei, 1987), as implemented in *pegas*, and the mean and standard deviation compared across estuaries. A Mann-Whitney test was used to test for a significant difference in mean Θ -values across estuaries and between ocean basins.

2.6 | Landscape genetics

Redundancy analysis (RDA), as implemented in *vegan* (Oksanen et al., 2013), was used to disentangle the influence of geographic distance and environmental variables and assess their interaction on observed patterns of genomic variation among samples from the Gulf. RDA was not carried out among samples from the Atlantic because of the limited number of geographic samples with enough individuals (3) and their uneven, limited geographic spread (see Figure 1). RDA is a constrained ordination method that extracts and summarizes components of variation in a multidimensional data set explained by a set of explanatory variables. It is a useful approach when using genomic data, that does not rely on equilibrium assumptions present in F_{ST} -based analyses (Forester et al., 2018). The R^2 value can be understood as the proportion of genomic variation explained by constraining variables, allowing for a comparison of the relative importance of these variables and their interaction. Here, the RDA was used to parse and visualize components of genomic variation (response variables) that are explained by geography and/or environment (constraining variables) and to identify alleles/loci driving any observed environmental pattern. To achieve this, two constraining matrices were generated, one describing spatial patterns and one

describing environmental differences among estuaries. For each, forward selection was used to identify the best model using R^2 as the stopping criterion (Oksanen et al., 2013). Geographic distance was measured as the approximate coastline distance between mouths of estuaries; distances were jittered for individual fish to account for individual variability in sample location. Data for 39 environmental variables for each of the included sample locations in the Gulf were obtained from the National Estuary Eutrophication Assessment database (NEEA, 2018). This data set was chosen as it includes all estuaries being evaluated and contains a broad range of variables describing the long-term hydrology; abiotic variables such as nutrients, temperature, salinity, pH, or suspended particles; and climate data, for example, humidity, precipitation, or wind speed. Because the goal was to enumerate environmental differences among estuaries, not necessarily to identify individual environmental pressures (for the latter environmental data collected concurrent to YOY residence in bays/estuaries would be more appropriate), variables were PCA-transformed, resulting in new synthetic variables summarizing environmental differences among estuaries. After identifying the best models describing components of genetic variation explained by spatial and environmental models alone, variance partitioning was used to compare the contribution of geographic distance and environmental differences in structuring observed genomic variation and disentangle whether geography or the environment plays a larger role in shaping genetic diversity. A full model, using geographic and PCA-transformed environmental variables, a partial model using geographic data conditioned on environmental variables, and a partial model using environmental variables conditioned on geographic data, were considered to partition the explainable variance into individual (geography or environment) and shared components (geography plus environment), using *vegan*. Significance of each component was tested using 1,000 permutations. The environmental model was then used to identify loci most strongly associated with environmental differences among estuaries. Alleles with a Mahalanobis distance >25 on the first two RDA-axes were flagged as most strongly associated with PCA-transformed environmental variables. The distribution of alleles (loci) having a Mahalanobis distance >25 across linkage groups was assessed using those loci that previously were incorporated into a linkage map (O'Leary et al., 2018).

All figures were generated using *ggplot2* (Wickham, 2009) and *UpsetR* (Conway et al., 2017). An Rmarkdown and corresponding rendered html-document containing reproducible code for the complete analysis and functions as a standalone extended presentation of methods and results can be accessed at https://github.com/sjoleary/SFL_PopGen.

3 | RESULTS

3.1 | F_{ST} -outlier analysis

The final filtered data included 316 individuals from six estuaries in the Atlantic and nine estuaries in the Gulf (Table 1, Figure 1). No

TABLE 2 Locus-by-locus AMOVA using only (A) neutral and (B) outlier loci indicating variance partitioning, F -statistic, and significance of each component

Source of variation	Percentage variation	Average F -statistic over all loci	p -Value
(A)			
Among oceans	4.1348	0.0414	<.0001*
Among estuaries within oceans	0.1551	0.0016	.0093*
Among individuals within estuaries	95.7101	0.0429	<.0001*
(B)			
Among oceans	32.7517	0.3275	<.0001*
Among estuaries within oceans	0.1332	0.0020	.6020
Among individuals within estuaries	67.1150	0.3289	<.0001*

Note: Significant values are in bold and marked with *.

F_{ST} -outlier loci were detected using either approach when Gulf or Atlantic individuals were assessed separately.

3.2 | Assessment of population structure and genetic diversity

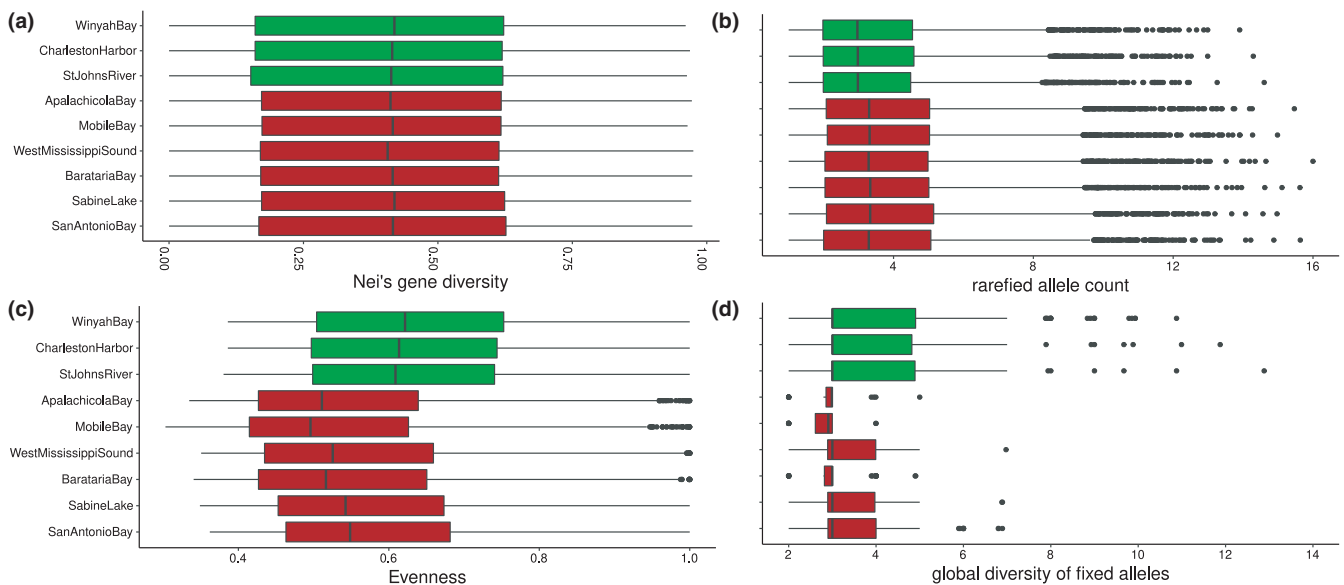
Hierarchical AMOVA implemented to test for heterogeneity between and within ocean basins revealed significant divergence ($p < .0001$) between the Atlantic and Gulf in both neutral and outlier loci (Table 2). The magnitude of F_{CT} differed between the two data sets: 0.0414 (neutral loci) and 0.3275 (outlier loci). Further, divergence in neutral loci among estuaries within ocean basins was significant ($F_{SC} = 0.0016$, $p = .0093$), whereas divergence in outlier loci was not ($F_{SC} = 0.0020$, $p = .6020$). Divergence among estuaries using neutral loci only was significant in the Gulf ($F_{ST} = 0.0014$, $p = .027$) but not in the Atlantic ($F_{ST} = 0.0026$, $p = .098$, Table 2). By contrast, all pairwise comparisons between estuaries in different ocean basins were significant ($p < .001$) for both neutral and outlier loci (Table 3). For neutral loci, comparison of San Antonio Bay (SA) and Sabine Lake (SL) was significant before (but not after) correction for multiple tests, while San Antonio Bay (SA) and Apalachicola Bay (AP) remained significantly different even after correction for multiple tests. All remaining comparisons between estuaries within ocean basins were nonsignificant for both neutral and outlier loci (Table 3).

The comparison of genomic diversity among estuaries revealed significant heterogeneity among estuaries for all three metrics of within-group genetic diversity: Nei's gene diversity ($Q_{[8]} = 158.21$, $p < .0001$), rarefied allele counts ($Q_{[8]} = 1,536.5$, $p < .0001$), and evenness ($Q_{[8]} = 3,842.3$, $p < .0001$, Figure 2). Nei's gene diversity and rarefied allele counts were significantly higher in estuaries in the Gulf for all 18 pairwise comparisons with estuaries in the Atlantic (Figure 2a,b). All comparisons of gene diversity between

TABLE 3 Comparison of pairwise F_{ST} (above the diagonal) and level of significance (below the diagonal) between all pairs of estuaries in the Gulf and Atlantic using (A) neutral and (B) outlier loci only

	SA	SL	BB	WMS	MB	AP	SJR	CH	WB
(A)									
SA	-	0.0020	0.0010	0.0014	0.0011	0.0016*	0.0466*	0.0423*	0.0443*
SL	0.0329	-	0.0010	0.0010	0.0004	0.0010	0.0452*	0.0411*	0.0427*
BB	0.2101	0.3244	-	0.0006	0.0007	0.0007	0.0462*	0.0421*	0.0439*
WMS	0.0644	0.3757	0.3804	-	0.0004	0.0005	0.0462*	0.0420*	0.0441*
MB	0.1253	0.9022	0.0876	0.6603	-	0.0006	0.0450*	0.0405*	0.0428*
AP	0.0040	0.2847	0.1480	0.5679	0.2009	-	0.0457*	0.0414*	0.0432*
SJR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-	0.0010	0.0011
CH	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.4436	-	0.0011
WB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6043	0.1469	-
(B)									
SA	-	-0.0011	0.0019	0.0046	-0.0002	0.0021	0.3399*	0.3419*	0.3055*
SL	0.8747	-	0.0006	0.0050	0.0017	0.0014	0.3419*	0.3445*	0.3063*
BB	0.4627	0.6881	-	0.0008	0.0017	0.0026	0.3480*	0.3436*	0.3101*
WMS	0.1503	0.1274	0.5995	-	0.0019	0.0035	0.3265*	0.3244*	0.2906*
MB	0.7789	0.3626	0.2657	0.2619	-	0.0015	0.3400*	0.3347*	0.3032*
AP	0.3948	0.5244	0.1920	0.1163	0.2855	-	0.3481*	0.3439*	0.3122*
SJR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-	-0.0077	0.0008
CH	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9928	-	-0.0003
WB	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.6027	0.6214	-

Note: Significant values are in bold and marked with *.

**FIGURE 2** Distribution of (a) Nei's gene diversity, (b) rarefied allele count, (c) evenness, and (d) global diversity of fixed alleles. Estuaries in the Gulf and Atlantic are represented in red and green, respectively

estuaries within basins were nonsignificant except for comparisons with Sabine Lake, which had significantly higher gene diversity than all other estuaries in the Gulf (Figure 2b). Similarly, all comparisons of rarefied allele counts between estuaries within basins were nonsignificant except for comparisons with Sabine Lake, which had significantly higher allele counts than all other estuaries in the Gulf, and

Mobile Bay which had significantly higher allele counts than West Mississippi Sound (Figure 2b). Evenness was significantly lower in estuaries in the Gulf for all pairwise comparisons with estuaries in the Atlantic (Figure 2c) and no significant differences were found within basins. A comprehensive table with all pairwise tests for all three metrics is found in Table S1. Finally, there was a higher frequency of

fixed alleles in estuaries in the Atlantic as compared with the Gulf (Figure 3). A comparison of the intersection of sets of fixed alleles indicated that the largest intersects were one-set intersects, that is, loci fixed in a single Atlantic estuary (44–75), a three-set intersect of loci fixed in all three Atlantic estuaries (73) and two-set intersects consisting of combinations of two Atlantic estuaries (Figure 3). In general, loci fixed in Gulf estuaries had low global diversity, while loci fixed in the Atlantic were a mixture of loci with low and high global diversity (Figure 2d).

3.3 | Tests of genome and locus neutrality

The genome-wide distribution of Tajima's D was determined for each estuary and compared with an equivalent simulated data set consisting of loci in neutral, drift-mutation equilibrium to test for patterns consistent with past demographic events. For each estuary, the observed genome-wide distributions of Tajima's D values were left-shifted relative to simulated equilibrium distributions and observed mean and median values of Tajima's D were significantly more negative than the simulated values for each estuary ($p < .001$; Table S2). This pattern is consistent with population expansion.

Locus-specific effects were assessed by testing for the significant deviation from mutation-drift equilibrium for each locus for individuals grouped by estuary. A total of 422 loci (10.2% of all loci examined) were significantly different from the expectation of equilibrium ($p < .05$) in at least one estuary; after accounting for multiple

comparisons 12 remained significant: 63 (5) were positive, 361 (9) were negative, and 167 (2) were positive in at least one estuary and negative in at least one other estuary. The number of significant positive loci ranged from 15 to 21 across estuaries in the Atlantic; after adjusting for multiple comparisons Winyah Bay and Charleston Harbor had one negative outlier each, while none were found in St John's River individuals, compared from 11 to 19 across estuaries in the Gulf (2–4 after adjusting p -values). The number of significant negative loci across estuaries in the two ocean basins differed markedly; estuaries in the Gulf averaged 91.7 significant negative loci (range = 80–108; 1–4 after adjustment), whereas estuaries in the Atlantic averaged 25.3 (range = 20–31); after adjusting for multiple comparisons, Charleston Harbor and St John's River each had one significant outlier while none were observed in Winyah Bay. A total of 155 loci with significant positive (52), or negative (101), or both (2) Tajima's D values in at least one estuary had been mapped previously and appeared to map randomly across all 24 linkage groups (Figure S1), suggesting no specific regions of the genome were in disequilibrium.

Tajima's D is derived from a comparison of the number of segregating sites and the average pairwise difference among all sequences, and to understand the mechanisms contributing to departure from mutation drift-equilibrium these two metrics were compared. Mean values of Watterson's estimator Θ_W (based on the number of segregating sites) were significantly higher and more variable in Gulf estuaries ($\Theta_W = 0.454$ – 0.602) as compared with Atlantic estuaries ($\Theta_W = 0.353$ – 0.370 , Table S3). By contrast, mean values for Θ_T (based

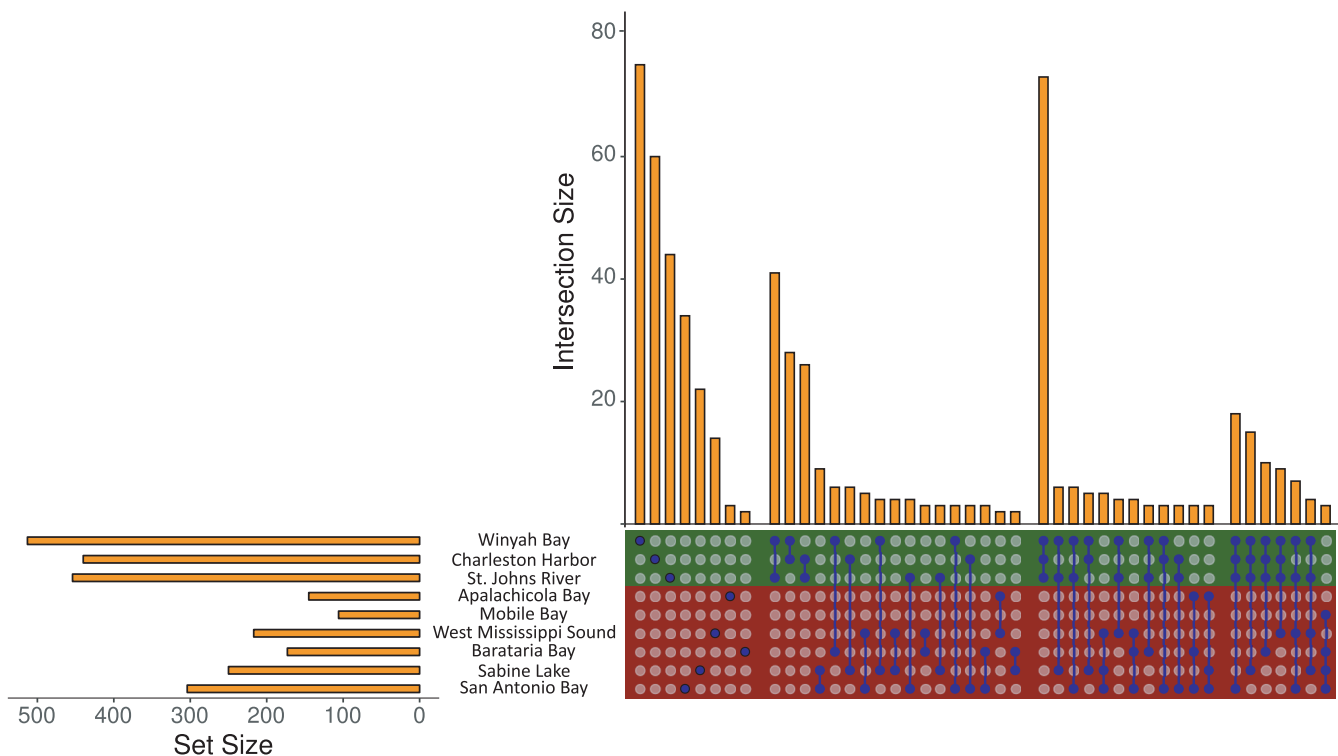


FIGURE 3 Comparison of fixed loci across sampled estuaries in the Gulf (red box) and Atlantic (green box). The set size (horizontal orange bars) indicates the total number of loci fixed in a given location, the intersection size (vertical orange bars) indicates the number of loci fixed only in a single location (single blue dot) or in two, three, or four locations (indicated by blue dots connected by line)

on pairwise differences among haplotypes) had a much smaller range ($\Theta_T = 0.0020\text{--}0.0021$, Table S4) across all estuaries. Sabine Lake had significantly higher Θ_T than all other bays, while Winyah Bay displayed the opposite pattern (Table S4). Mann–Whitney tests indicated that both Θ estimates were significantly higher in the Gulf compared with the Atlantic ($P(\Theta_{\text{Gulf}}) = 0.0282$, $P(\Theta_T) = 0.0282$).

3.4 | Landscape genomics

The RDA framework was used to test for the effects of geographic distance and environmental differences among estuaries in the Gulf on genomic diversity, to identify loci strongly associated with environmental differences, and to partition the components of variance associated with distance, environment, and their interactions. The selected model for geographic distance included the third polynomial of coastal distance, indicative of a pattern of nonlinear isolation by distance. The selected model for environmental differences among estuaries included principal components eight and four. Loading plots for both principal components (Figure 4) indicated that parameters with largest impact include variables related to freshwater inflow (e.g., total suspended sediment, total freshwater volume), tidal influence (e.g., tidal exchange, estuary depth, tide volume, tide ratio), factors determining differences in salinity within and among estuaries (daily evaporation, freshwater volume, percent mixed water, percent seawater), and amount of available wetland habitat.

After determining the best models describing spatial and environmental patterns across estuaries, variance partitioning was performed to evaluate relative impacts of spatial patterns compared with environmental differences. The full model that included both the selected spatial and environmental model was significant ($p = .023$), though it explained only a small proportion of variance in the total genomic data set (1.3%, adjusted $R^2 = .00023$). The largest component of variance was explained by environmental variables and shared effects (adjusted $R^2 = .00026$, $p = .005$), while the component of spatial, environmental, and shared effects was marginally lower (adjusted $R^2 = .00024$, $p = .029$). The smallest significant component of the full model was spatial and shared effects (adjusted $R^2 = .00014$, $p = .020$, Table 4).

Clustering of individuals was compared for the full model and environmental model only (Figure 5). A biplot of the full RDA model (Figure 5a) revealed a complex pattern with individuals from the same estuary forming clusters, but with placement of individual estuaries along individual RDA axes not always corresponding to geographic proximity. San Antonio Bay and Mobile Bay and West Mississippi Sound formed two clusters separated from other samples along RDA 1 (Quadrant II, III) driven by the spatial matrix and PC 8. Individuals from the remaining estuaries were resolved by RDA 2, with Apalachicola Bay and Barataria Bay forming two distinct clusters in Quadrant I and East Mississippi Sound, Galveston Bay and Sabine Lake/Matagorda Bay individuals forming clusters in Quadrant IV. The environmental model clustered individuals in a very similar pattern (Figure 5b), the difference being how the clusters lie relative

to each other in Quadrant IV. Sabine Lake and Matagorda Bay individuals clustered separately, and Sabine Lake and Galveston Bay individuals were closer to the Mobile Bay/West Mississippi Sound cluster, while Matagorda remained in the lower right-hand corner of Quadrant IV. Alleles at 384 loci (9.3%) were flagged as strongly associated with both RDA axis of the environmental model (Mahalanobis Distance >25). Of these, 123 had been previously mapped on the linkage map and appeared to be randomly distributed across all chromosomes (Figure S1).

4 | DISCUSSION

Here, reduced representation sequencing was used to identify patterns of contemporary population structure of southern flounder sampled in the U.S. Atlantic and Gulf of Mexico. Significant divergence was observed between the Atlantic and Gulf. While A_{MOVA} based on neutral markers indicated significant heterogeneity among estuaries within the Gulf, most pairwise comparisons of sampled estuaries within ocean basins were not significant. Overall, Gulf estuaries exhibited significantly higher levels of allelic diversity as compared with the Atlantic estuaries, though no explicit spatial patterns of within-sample diversity were detected within the Gulf and estuaries in both basins exhibited mutation-drift disequilibrium at genome-wide scales. Landscape genomic approaches revealed a pattern of nonlinear isolation by distance across the northern Gulf and indicated that environmental variables summarizing differences among individual estuaries explain a larger portion of genomic variation than geographic distance.

Significant divergence between Atlantic and Gulf populations of southern flounder aligns with a well-documented biogeographic break associated with southern peninsular Florida (Neigel, 2009). Genetic discontinuities associated with southern Florida have been observed in a number of nearshore and estuarine species (Gold & Richardson, 1998) and are likely related to a lack of appropriate estuarine habitat. Consistent with this, southern flounder show a gap in their distribution in southern Florida, approximately between Tampa and Cape Canaveral (Ginsburg, 1952). Given adult movement ecology in the Gulf and Atlantic, larval dispersal, and the results here, it seems highly unlikely that the Atlantic and Gulf experience biologically important levels of contemporary gene flow.

F_{ST} -outlier loci were detected when analyzing Gulf and Atlantic samples together. By contrast, no outlier loci were detected when Atlantic and Gulf data sets were analyzed separately. F_{ST} -outliers are generally considered to be under positive selection, though historical demographic processes may create similar patterns (Lotterhos & Whitlock, 2014). Here, analysis of Tajima's D indicated a genome-wide departure from mutation-drift equilibrium consistent with population expansion, rather than departures at specific loci only, which would be consistent with directional selection. This mirrors other studies that have detected postglacial expansion in marine species from the Gulf and U.S. South Atlantic (Marko & Hart, 2012; Portnoy et al., 2014). Furthermore, loci flagged as F_{ST} -outliers or with

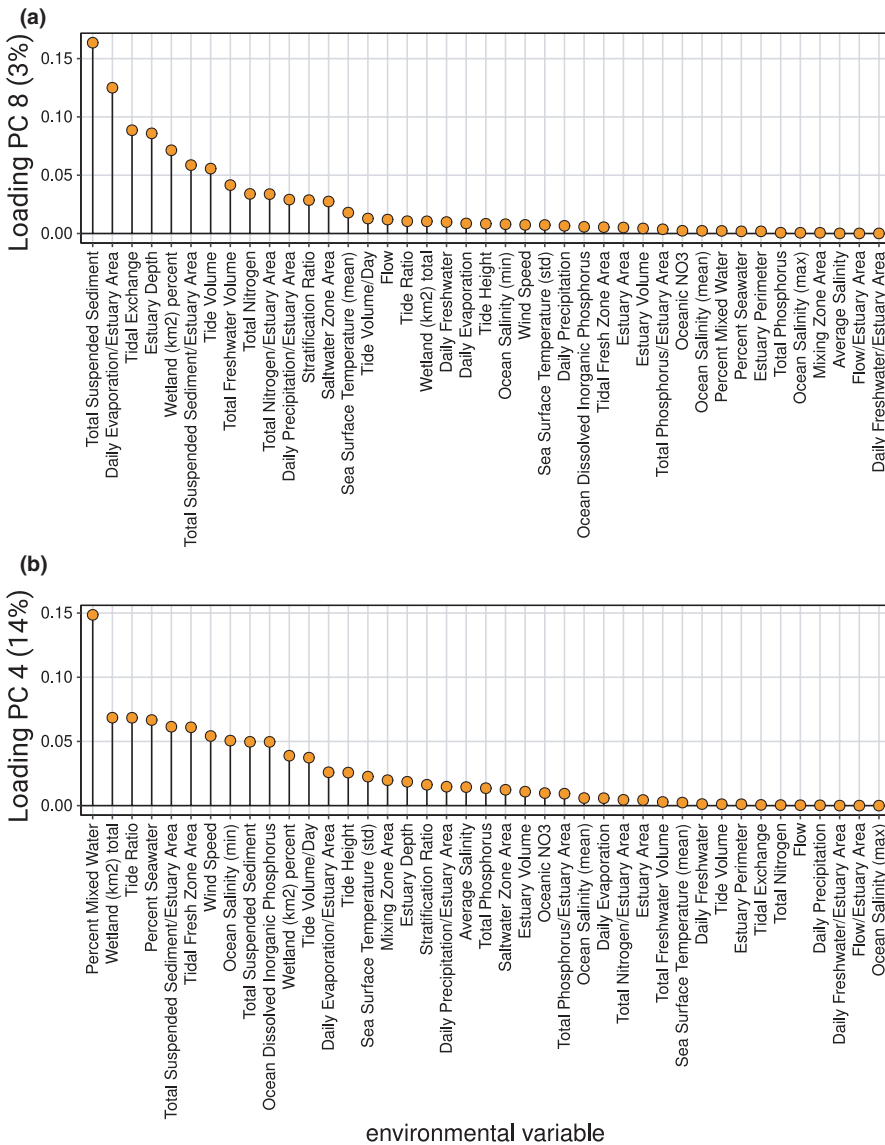


FIGURE 4 Loading plots for principal component analysis of environmental variables describing sampled estuaries for (a) principal component 8 and (b) principal component 4 which explain 3% and 14% of environmental differences among estuaries, respectively

TABLE 4 Partitioning of variance explained by costal distance (xy), environmental variables (env) and shared effects (shared) due to interaction of distance and environment

Partition	Variance (adjusted R^2)	p -Value
residuals	0.99976	n/a
env + shared	0.00026	.005*
env + xy + shared	0.00024	.029*
shared	0.00016	n/a
xy + shared	0.00014	.020*
env	0.00010	.200
xy	-0.00002	.560

Note: Significant values in bold and marked with *.

significantly negative Tajima's D (both taken to indicate directional selection) were spread across the genome rather than clustered together, the latter expected if strong, locus-specific directional selection was present (Nielsen, 2001). Simulation studies have shown

that differentiation-based outlier tests are not robust when applied to nonequilibrium demographic scenarios, specifically isolation-by-distance and range expansions, and may suffer increased rates of false positives (Forester et al., 2018; Hoban et al., 2016; Lotterhos & Whitlock, 2014). While directional selection operating at an ocean basin scale cannot be ruled out, the agreement in pattern between putatively neutral loci and outlier loci along with evidence of recent expansion, suggest that F_{ST} outlier-loci identified in this study are drift outliers.

Traditional F_{ST} -based approaches assume that microevolutionary forces are in equilibrium and simple demographic models, such as island or stepping stone models, are sufficient to explain observed patterns of genetic heterogeneity (Holsinger & Weir, 2009). When these assumptions are violated, which is the case for many marine species (Waples, 1998), spatially explicit analyses are preferable (Manel et al., 2003). For southern flounder within the Gulf, pairwise estimates of F_{ST} between estuaries were generally not significant, despite global tests indicating significant genetic heterogeneity across the Gulf. By contrast, RDA indicated spatial autocorrelation

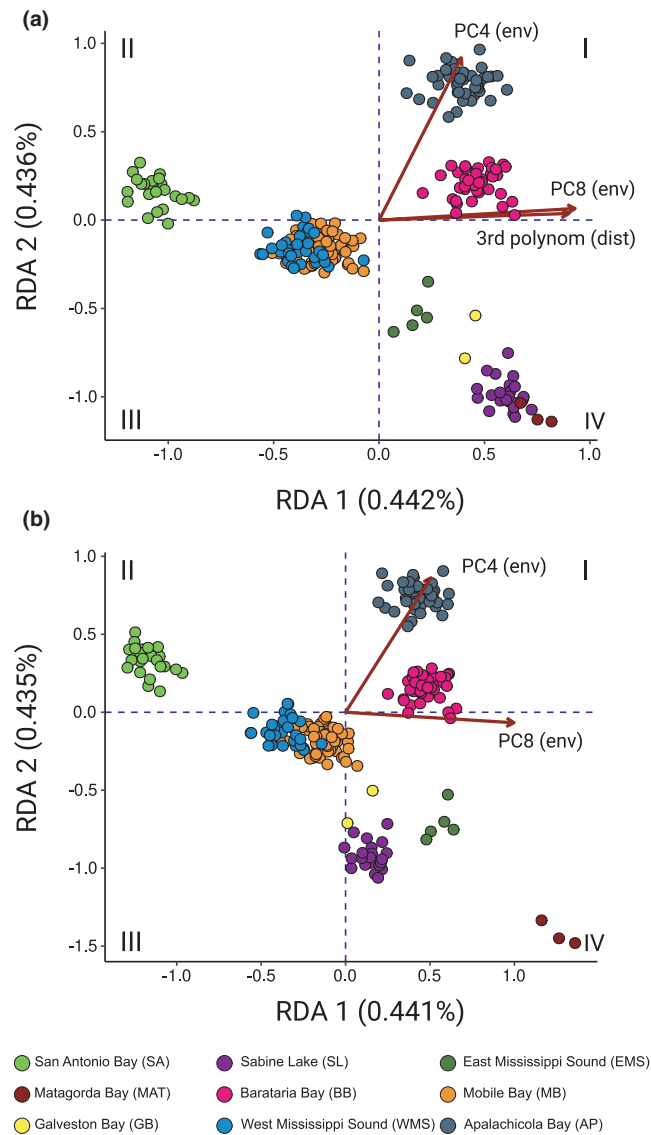


FIGURE 5 Biplot of redundancy analysis using (a) a full model using explanatory matrices selected for the best spatial model (3rd degree polynomial of coastal distance) and environmental model (PC8, PC4) and (b) the environmental model alone. Individuals sampled in the Gulf (colored circles) are plotted according to their component loadings calculated as weighted average scores. The full model explains 1.3% of variance, RDA1 and 2 explain 33.95% (0.442%) and 33.40% (0.436%) of constrained (total) variance, respectively. The environmental model explains 0.8% of variance, RDA1 and 2 explain 50.31% (0.441%) and 49.69% (0.435%) of constrained (total) variance. Quadrants I–IV are labeled for easier reference in the text

and a pattern of nonlinear isolation by distance among estuaries in the Gulf. Further, environmental similarities (or differences) among habitats explained a significant component of total genomic variation that was larger than geographic distance, even though F_{ST} outlier tests found no significant departure from neutral expectations in the Gulf. Environmental conditions may change allele frequencies in concert with or in opposition to drift processes (Kawecki & Ebert, 2004), and the latter seems to have occurred in southern

flounder. For example, individuals from estuaries in the western Gulf do not cluster together when constraining the variation to the component explained by the full RDA model (Figure 5a). Furthermore, the position of the Sabine Lake cluster relative to other estuaries changes when constraining the variance to the component explained by the environmental model (Figure 5b). The shift in how individuals from estuaries in Quadrant IV cluster relative to each other, depending on whether or not spatial signal is explicitly accounted for, has real world applications. Currently, southern flounder are part of a stock-enhancement program in the western Gulf where captive wild-caught adults are spawned in a hatchery and YOY head-started in outdoor ponds. The extent to which YOY should be returned to the estuaries from which their parents were caught or whether they can be efficiently stocked in neighboring estuaries has remained an open question (Kaiser et al., 2012; Miller et al., 2010). The results present here indicate that it is important to consider environmental differences among estuaries as well as geographic proximity.

While the full RDA model only explained a small percentage of total variation (approx. 1%), with the majority attributable to environment, the results are not that dissimilar from other studies. Bay et al. (2017) presented a review of genomic studies and found that between 0.002% and 4.6% of loci, across studies, appeared to be influenced by environment. Furthermore, gene flow across estuaries combined with selection operating on multiple genes but of small effect can result in a swamping of the signal (Yeaman, 2015). A similar result was seen in a study of the congeneric summer flounder, *Paralichthys dentatus*, in the Atlantic, where environmental factors appear to drive small but important differences in genetic diversity despite near-panmixia and no evidence of significant population structure across most of the genome (Hoey & Pinsky, 2018).

For southern flounder, spawning takes place offshore from nursery areas and prevailing currents within and between years determine the strength and sources of larval influx, though habitat quality and availability for settling may ultimately dictate recruitment (Burke et al., 1991; Miller et al., 1991). Consistent with this, environmental variables of importance identified using the RDA framework included tidal influence and freshwater input, factors impacting differences in salinity among and within estuaries, as well as available wetland habitat. These results parallel studies that have identified environmental variables, including estuary depth and slope, tidal height, habitat type, proximity to inlet, temperature, salinity, turbidity, and levels of dissolved oxygen, as determinates of the density of juvenile southern flounder within estuaries (Glass et al., 2008; Nañez-James et al., 2009). Further, these studies indicate that preferences for combinations of conditions seem to vary across estuaries. For example, juvenile southern flounder in estuaries off North Carolina were usually sampled in shallow water with low salinity, high dissolved oxygen, and muddy bottom substrates, and found far from the estuary inlet (Burke et al., 1991; Powell & Schwartz, 1977); while in Aransas Bay (Texas), juvenile abundance was highest in vegetated, sandy areas with higher salinities and located near estuary inlets, with low abundances in muddy bottoms habitat (Nañez-James et al., 2009). Additionally, estuaries are a mosaic of microhabitats and

juvenile southern flounder exhibit ontogenetic shifts in microhabitat use (Amorim et al., 2018; Furey & Rooker, 2013; Miller et al., 1991). In the Gulf, juvenile southern flounder initially settle into structurally complex habitats like seagrass beds and marsh edges, transitioning to habitats characterized by sandy or muddy substrates toward the end of their juvenile stages (Furey & Rooker, 2013; Glass et al., 2008; Nañez-James et al., 2009). Access to appropriate habitat during development may strongly affect survival of juveniles at a time when mortality rates are high and the availability and types of habitat differ among estuaries (Amorim et al., 2018; Burke et al., 1991). This sets up a complex scenario in which gene flow mediated by larval, and perhaps adult, dispersal is related to distance, while environmental and physiochemical conditions that vary among and within estuaries dictate survival of recruits (Burke et al., 1998). This results in components of genetic diversity shaped initially by dispersal (migration and drift), and subsequently refined by characteristics of the local habitat experienced by juveniles (selection) in such a way that there may be conflicting signal within the genome. Estuaries by their nature are dynamic habitats that vary environmentally across years and such a dynamic could lead to an association between genetic variation and interannual environmental and climate variability but would require temporally explicit samples across cohorts to detect, which was not possible in this study.

Additionally, environmental heterogeneity and ontogenetic shifts in habitat use may facilitate the maintenance of diversity (balancing selection), rather than favor specific phenotypes (directional/purifying selection), when survivorship is determined by genotype-environment matches at the microhabitat scale (Bernatchez et al., 2019; Marshall et al., 2010). The complex interactions of competing evolutionary forces, along with diversifying or weak polygenic selection can result in genetic diversity characterized by the presence of rare alleles (Kawecki & Ebert, 2004; Star & Spencer, 2013), as seen in southern flounder samples in the Gulf. Consistent with this idea, loci with large Mahalanobis distance (>25) were spread throughout the genome, rather than grouped, and there were no significant F_{ST} -outlier loci in the Gulf indicative of strong directional selection. Furthermore, patterns of reduced genetic diversity in the Atlantic relative to the Gulf are largely attributable to a reduction in the prevalence of rare alleles and could be due to less environmental heterogeneity, but also stronger genetic drift and/or a smaller source population postglacial expansion (Allendorf, 1986). While the exact mechanisms cannot be inferred due to limited sampling in the Atlantic, estimates of Θ_W and Θ_π were both smaller in Atlantic estuaries than the Gulf estuaries, suggesting that reduced long-term effective population size in the Atlantic may at least partially explain the pattern. The presence of multiple forces shaping genomic diversity, to an extent in opposing ways, likely contributes to distinct signals parsed in this study each only explaining small though significant proportions of total variance (Yeaman, 2015).

Understanding the interplay of microevolutionary processes has important implications for marine species, especially for those like southern flounder that are the focus of commercial/recreation fisheries and stock augmentation programs. A pattern of high gene flow

and isolation by distance alone would suggest that the geographic origin of fish might matter only on large spatial scales for stocking purposes. By contrast, the finding that environmental differences among estuaries explains total genomic variation better than relative geographic position indicates that best management strategies for southern flounder should include a focus on the preservation of a diversity of habitats that can be used to complete early life stages. Ecosystem-based management approaches already emphasize the preservation of critical habitat to sustain populations but often argue from a standpoint of the importance of certain specific habitats (Rosenberg et al., 2000). In this study, environmental-genome associations were related to relative differences (or similarities) in the environment among estuaries, and environmental characteristics of particular estuaries explained a significant but small proportion (approx. 1%) of genomic variation among estuaries. The observations that the environmental conditions that contribute to survival and successful reproduction (fitness) vary in space suggests that not only the availability of specific habitats but also the diversity of available habitat types is important. Further, in southern flounder selection appears to be complex, involving many loci spread throughout the genome, and the results add to a growing body of research that demonstrates the importance of considering models of gene-environment interactions without equilibrium assumptions, rather than only screening for F_{ST} -based outlier loci indicative of strong directional selection (Forester et al., 2018; Hoban et al., 2016).

ACKNOWLEDGMENTS

The authors thank the individuals (M. Drymon, B. Falterman, E. Saillant, M. Fisher, T. Darden) and agencies (Florida Fish and Wildlife Conservation Commission, Texas Parks and Wildlife Department, and South Carolina Department of Natural Resources) who assisted in procuring tissues. They also thank A. Fields and members of the marine genomics working group at TAMUCC for many helpful discussions. In addition, the authors thank John Gold for comments, discussion, and input into earlier drafts of this manuscript. Research was funded through the State Wildlife Grant Program-CFDA# 15.634. This article is publication number 28 of the Marine Genomics Laboratory at Texas A&M University - Corpus Christi and number 121 in the series Genetic Studies in Fishes.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

Shannon J. O'Leary: Conceptualization (supporting); Data curation (lead); Formal analysis (lead); Methodology (lead); Project administration (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (equal). **David S. Portnoy:** Conceptualization (lead); Formal analysis (supporting); Investigation (supporting); Methodology (supporting); Project administration (supporting); Resources (lead); Writing-review & editing (equal). **Christopher M. Hollenbeck:** Formal analysis (supporting); Methodology (supporting); Writing-review & editing (equal). **Robert R. Vega:** Conceptualization (supporting); Funding (supporting).

DATA AVAILABILITY STATEMENT

Raw reads are submitted at the short reads archive (BioProject PRJNA754841). Processed data sets underlying the analysis, custom scripts, and R markdown files with reproducible code are available at https://github.com/sjoleary/SFL_PopGen.

ORCID

Shannon J. O'Leary  <https://orcid.org/0000-0001-9775-9846>

REFERENCES

- Allendorf, F. W. (1986). Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, 5(2), 181–190. <https://doi.org/10.1002/zoo.1430050212>
- Amorim, E., Ramos, S., Elliott, M., & Bordalo, A. A. (2018). Dynamic habitat use of an estuarine nursery seascape: Ontogenetic shifts in habitat suitability of the European flounder (*Platichthys flesus*). *Journal of Experimental Marine Biology and Ecology*, 506, 49–60. <https://doi.org/10.1016/j.jembe.2018.05.011>
- Anderson, J. D., Karel, W. J., & Mione, A. C. S. (2012). Population structure and evolutionary history of southern flounder in the Gulf of Mexico and Western Atlantic Ocean. *Transactions of the American Fisheries Society*, 141(1), 46–55. <https://doi.org/10.1080/00028487.2011.651070>
- Baetscher, D. S., Clemento, A. J., Ng, T. C., Anderson, E. C., & Garza, J. C. (2018). Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, 18(2), 296–305. <https://doi.org/10.1111/1755-0998.12737>
- Bailey, K. M., Nakata, H., & Van Der Veer, H. W. (2005). The planktonic stages of flatfishes: Physical and biological interactions in transport processes. In *Flatfishes: Biology and exploitation* (pp. 94–119).
- Bay, R. A., Rose, N., Barrett, R., Bernatchez, L., Ghalambor, C. K., Lasky, J. R., Brem, R. B., Palumbi, S. R., & Ralph, P. (2017). Predicting responses to contemporary environmental change using evolutionary response architectures. *American Naturalist*, 189(5), 463–473. <https://doi.org/10.1086/691233>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57(1), 289–300. <https://doi.org/10.2307/2346101>
- Bernatchez, L. (2016). On the maintenance of genetic variation and adaptation to environmental change: Considerations from population genomics in fishes. *Journal of Fish Biology*, 89(6), 2519–2556. <https://doi.org/10.1111/jfb.13145>
- Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., Maes, G. E., Martinsohn, J. T., Miller, K. M., Naish, K. A., Ovenden, J. R., Primmer, C. R., Young Suk, H. O., Therkildsen, N. O., & Withler, R. E. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology and Evolution*, 32(9), 665–680. <https://doi.org/10.1016/j.tree.2017.06.010>
- Bernatchez, S., Xuereb, A., Laporte, M., Benestan, L., Steeves, R., Laflamme, M., & Mallet, M. A. (2019). Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. *Evolutionary Applications*, 12(3), 587–609. <https://doi.org/10.1111/eva.12741>
- Blandon, L. R., Ward, R., King, T. L., Karel, W. J., & Monaghan, J. P. (2001). Preliminary genetic population structure of southern flounder, *Paralichthys lethostigma*, along the Atlantic Coast and Gulf of Mexico. *Fishery Bulletin*, 99(4), 671–678.
- Burke, J. S., Miller, J. M., & Hoss, D. E. (1991). Immigration and settlement pattern of *Paralichthys dentatus* and *P. lethostigma* in an estuarine nursery ground, North Carolina, U.S.A. *Netherlands Journal of Sea Research*, 27(3–4), 393–405. [https://doi.org/10.1016/0077-7579\(91\)90041-X](https://doi.org/10.1016/0077-7579(91)90041-X)
- Burke, J. S., Ueno, M., Tanaka, Y., Walsh, H., Maeda, T., Kinoshita, I., Tanaka, M. (1998). The influence of environmental factors on early life history patterns of flounders. *Journal of Sea Research*, 40, 19–32. [https://doi.org/10.1016/S1385-1101\(98\)00014-8](https://doi.org/10.1016/S1385-1101(98)00014-8)
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks : Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, 1(3), 171–182. <https://doi.org/10.1534/g3.111.000240>
- Clarke, L. M., Munch, S. B., Thorrold, S. R., & Conover, D. O. (2010). High connectivity among locally adapted populations of a marine fish (*Menidia menidia*). *Ecology*, 91(12), 3526–3537. <https://doi.org/10.1890/09-0548.1>
- Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics*, 33(18), 2938–2940. <https://doi.org/10.1093/bioinformatics/btx364>
- Corey, M. M., Leaf, R. T., Brown-Peterson, N. J., Peterson, M. S., Clardy, S. D., & Dippold, D. A. (2017). Growth and spawning dynamics of southern flounder in the north-central Gulf of Mexico. *Marine and Coastal Fisheries*, 9(1), 231–243. <https://doi.org/10.1080/19425120.2017.1290722>
- Cormack, R. M., Hartl, D. L., & Clark, A. G. (1990). Principles of population genetics. *Biometrics*, 46(2), 546. <https://doi.org/10.2307/2531471>
- Craig, J. K., Smith, W. E., Scharf, F. S., & Monaghan, J. P. (2015). Estuarine residency and migration of southern flounder inferred from conventional tag returns at multiple spatial scales. *Marine and Coastal Fisheries*, 7(1), 450–463. <https://doi.org/10.1080/19425120.2015.1079578>
- Dahlgren, C. P., & Eggleston, D. B. (2000). Ecological processes underlying ontogenetic habitat shifts in a coral reef fish. *Ecology*, 81(8), 2227–2240. [https://doi.org/10.1890/0012-9658\(2000\)081\[2227:EPUOH\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[2227:EPUOH]2.0.CO;2)
- Daniels, H. V., Berlinsky, D. L., Hodson, R. G., & Sullivan, C. V. (1996). Effects of stocking density, salinity, and light intensity on growth and survival of Southern flounder *Paralichthys lethostigma* larvae. *Journal of the World Aquaculture Society*, 27(2), 153–159. <https://doi.org/10.1111/j.1749-7345.1996.tb00264.x>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Felsenstein, J. (1976). The theoretical population genetics of variable selection and migration. *Annual Review of Genetics*, 10(1), 253–280. <https://doi.org/10.1146/annurev.ge.10.120176.001345>
- Fischer, A. J., & Thompson, B. A. (2004). The age and growth of southern flounder, *Paralichthys lethostigma*, from Louisiana estuarine and offshore waters. *Bulletin of Marine Science*, 75(1), 63–77.
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), 977–993. <https://doi.org/10.1534/genetics.108.092221>
- Flounder Technical Task Force (2015). *Management profile for the gulf and Southern flounder fishery in the Gulf of Mexico* (pp. 247). Gulf States Marine Fisheries Commission.
- Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, 27(9), 2215–2233. <https://doi.org/10.1111/mec.14584>
- Froeschke, B. F., Sterba-Boatwright, B., & Stunz, G. W. (2011). Assessing southern flounder (*Paralichthys lethostigma*) long-term population trends in the northern Gulf of Mexico using time series analyses. *Fisheries Research*, 108(2–3), 291–298. <https://doi.org/10.1016/j.fishres.2010.12.023>
- Furey, N. B., & Rooker, J. R. (2013). Spatial and temporal shifts in suitable habitat of juvenile southern flounder (*Paralichthys lethostigma*).

- Journal of Sea Research*, 76, 161–169. <https://doi.org/10.1016/j.seares.2012.08.007>
- Gagnaire, P. A., & Gaggiotti, O. E. (2016). Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. *Current Zoology*, 62(6), 603–616. <https://doi.org/10.1093/cz/zow088>
- Ginsburg, I. (1952). Flounders of the genus *Paralichthys* and related genera in American waters. *Fishery Bulletin*, 52(71), 267–351.
- Glass, L. A., Rooker, J. R., Kraus, R. T., & Holt, G. J. (2008). Distribution, condition, and growth of newly settled southern flounder (*Paralichthys lethostigma*) in the Galveston Bay Estuary, TX. *Journal of Sea Research*, 59(4), 259–268. <https://doi.org/10.1016/j.seares.2008.02.006>
- Gold, J. R., & Richardson, L. R. (1998). Mitochondrial DNA diversification and population structure in fishes from the Gulf of Mexico and western Atlantic. *Journal of Heredity*, 89(5), 404–414. <https://doi.org/10.1093/jhered/89.5.404>
- Hedgecock, D., & Pudovkin, A. I. (2011). Sweepstakes reproductive success in highly fecund marine fish and shellfish: A review and commentary. *Bulletin of Marine Science*, 87, 971–1002. <https://doi.org/10.5343/bms.2010.1051>
- Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*, 188(4), 379–397. <https://doi.org/10.1086/688018>
- Hoey, J. A., & Pinsky, M. L. (2018). Genomic signatures of environmental selection despite near-panmixia in summer flounder. *Evolutionary Applications*, 11(9), 1732–1747. <https://doi.org/10.1111/eva.12676>
- Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting FST. *Nature Reviews Genetics*, 10(9), 639–650. <https://doi.org/10.1038/nrg2611>
- Hudson, R. R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics*, 18(2), 337–338. <https://doi.org/10.1093/bioinformatics/18.2.337>
- Kaiser, J., Faulk, C., Williamson, E., & Holt, G. J. (2012). Natural spawning and larviculture of southern flounder *Paralichthys lethostigma*. *World Aquaculture*, 43(1), 48–54.
- 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, e281. <https://doi.org/10.7717/peerj.281>
- Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12), 1225–1241. <https://doi.org/10.1111/j.1461-0248.2004.00684.x>
- Lewontin, R. C., & Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74(1), 175–195. <https://doi.org/10.1093/genetics/74.1.175>
- Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, 23(9), 2178–2192. <https://doi.org/10.1111/mec.12725>
- Luikart, G., & Cornuet, J. M. (1998). Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*, 12(1), 228–237. <https://doi.org/10.1046/j.1523-1739.1998.96388.x>
- Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, 4(12), 981–994. <https://doi.org/10.1038/nrg1226>
- Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, 18(4), 189–197. [https://doi.org/10.1016/S0169-5347\(03\)00008-9](https://doi.org/10.1016/S0169-5347(03)00008-9)
- Marko, P. B., & Hart, M. W. (2012). Retrospective coalescent methods and the reconstruction of metapopulation histories in the sea. *Evolutionary Ecology*, 26(2), 291–315. <https://doi.org/10.1007/s10668-011-9467-9>
- Marshall, D. J., Monro, K., Bode, M., Keough, M. J., & Swearer, S. (2010). Phenotype-environment mismatches reduce connectivity in the sea. *Ecology Letters*, 13(1), 128–140. <https://doi.org/10.1111/j.1461-0248.2009.01408.x>
- Midway, S. R., Cadrin, S. X., & Scharf, F. S. (2014). Southern flounder (*Paralichthys lethostigma*) stock structure inferred from otolith shape analysis. *Fishery Bulletin*, 112(4), 326–338. <https://doi.org/10.7755/FB.112.4.9>
- Midway, S. R., Wagner, T., Arnott, S. A., Biondo, P., Martinez-Andrade, F., & Wadsworth, T. F. (2015). Spatial and temporal variability in growth of southern flounder (*Paralichthys lethostigma*). *Fisheries Research*, 167, 323–332. <https://doi.org/10.1016/j.fishres.2015.03.009>
- Miller, J. M., Burke, J. S., & Fitzhugh, G. R. (1991). Early life history patterns of Atlantic North American flatfish: Likely (and unlikely) factors controlling recruitment. *Netherlands Journal of Sea Research*, 27(3–4), 261–275. [https://doi.org/10.1016/0077-7579\(91\)90029-Z](https://doi.org/10.1016/0077-7579(91)90029-Z)
- Miller, J. M., Vega, R., & Yamashita, Y. (2010). Stock enhancement of southern and summer flounder. In *Practical flatfish culture and stock enhancement* (pp. 205–215).
- Monaghan, J. (1992). *Tagging studies of southern flounder (Paralichthys lethostigma) and gulf flounder (Paralichthys albigutta) in North Carolina*. Completion Report Project F-29, Study B.
- Morin, P. A., Martien, K. K., & Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, 9(1), 66–73. <https://doi.org/10.1111/j.1755-0998.2008.02392.x>
- Nañez-James, S. E., Stunz, G. W., & Holt, S. A. (2009). Habitat use patterns of newly settled southern flounder, *paralichthys lethostigma*, in aransas-copano bay, Texas. *Estuaries and Coasts*, 32(2), 350–359. <https://doi.org/10.1007/s12237-008-9107-y>
- NEAA (2018). *National Estuary Eutrophication Assessment*. <http://ian.umces.edu/nea/siteinformation.php>
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, 70(12), 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321>
- Nei, M. (1987). Molecular evolutionary genetics. *Molecular Evolutionary Genetics*, 512. <https://doi.org/10.7312/nei-92038>
- Neigel, J. E. (2009). Population genetics and biogeography of the Gulf of Mexico. In *Gulf of Mexico: Origins, Waters and Biota* (Vol. 1, pp. 1353–1369).
- Nielsen, R. (2001). Statistical tests of selective neutrality in the age of genomics. *Heredity*, 86(6), 641–647. <https://doi.org/10.1046/j.1365-2540.2001.00895.x>
- Nims, M. K., & Walther, B. D. (2014). Contingents of Southern Flounder from subtropical estuaries revealed by otolith chemistry. *Transactions of the American Fisheries Society*, 143(3), 721–731. <https://doi.org/10.1080/00028487.2014.892535>
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Wagner, H., & Oksanen, M. J. (2013). *Package 'vegan'*. *Community ecology package* (Vol. 2, pp. 1–295).
- O'Leary, S. J., Hollenbeck, C. M., Vega, R. R., Gold, J. R., & Portnoy, D. S. (2018). Genetic mapping and comparative genomics to inform restoration enhancement and culture of southern flounder, *Paralichthys lethostigma*. *BMC Genomics*, 19(1), 163. <https://doi.org/10.1186/s12864-018-4541-0>
- 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>

- Portnoy, D. S., Hollenbeck, C. M., Belcher, C. N., Driggers, W. B., Frazier, B. S., Gelsleichter, J., Grubbs, R. D., & Gold, J. R. (2014). Contemporary population structure and post-glacial genetic demography in a migratory marine species, the blacknose shark, *Carcharhinus acronotus*. *Molecular Ecology*, 23(22), 5480–5495. <https://doi.org/10.1111/mec.12954>
- Portnoy, D. S., Puritz, J. B., Hollenbeck, C. M., Gelsleichter, J., Chapman, D., & Gold, J. R. (2015). Selection and sex-biased dispersal: The influence of philopatry on adaptive variation. *PeerJ*, 1–20. <https://doi.org/10.7287/peerj.preprints.1300v1>
- Powell, A. B., & Schwartz, F. J. (1977). Distribution of Paralichthid Flounders (Bothidae: Paralichthys) in North Carolina Estuaries. *Chesapeake Science*, 18(4), 334. <https://doi.org/10.2307/1350587>
- Prince, D. J., O'Rourke, S. M., Thompson, T. Q., Ali, O. A., Lyman, H. S., Saglam, I. K., Hotaling, T. J., Spidle, A. P., & Miller, M. R. (2017). The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. *Science Advances*, 3(8), e1603198. <https://doi.org/10.1126/sciadv.1603198>
- Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent : A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, 2, e431. <https://doi.org/10.7717/peerj.431>
- Rosenberg, A., Bigford, T. E., Leathery, S., Hill, R. L., & Bickers, K. (2000). Ecosystem approaches to fishery management through essential fish habitat. *Bulletin of Marine Science*, 66, 535–542.
- Star, B., & Spencer, H. G. (2013). Effects of genetic drift and gene flow on the selective maintenance of genetic variation. *Genetics*, 194(1), 235–244. <https://doi.org/10.1534/genetics.113.149781>
- Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, 25(10), 2144–2164. <https://doi.org/10.1111/mec.13606>
- Wang, V. H., McCartney, M. C. A., & Scharf, F. S. (2015). Population genetic structure of southern flounder inferred from multilocus DNA profiles. *Marine and Coastal Fisheries*, 7(1), 220–232. <https://doi.org/10.1080/19425120.2015.1037473>
- Wang, V. H., Wilson White, J., Arnott, S. A., & Scharf, F. S. (2018). Population connectivity of southern flounder in the US South Atlantic revealed by otolith chemical analysis. *Marine Ecology Progress Series*, 596, 165–179. <https://doi.org/10.3354/meps12576>
- Waples, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, 89(5), 438–450. <https://doi.org/10.1093/jhered/89.5.438>
- Watanabe, W. O., Woolridge, C. A., & Daniels, H. V. (2006). Progress toward year-round spawning of southern flounder broodstock by manipulation of photoperiod and temperature. *Journal of the World Aquaculture Society*, 37(3), 256–272. <https://doi.org/10.1111/j.1749-7345.2006.00036.x>
- Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7(2), 256–276. [https://doi.org/10.1016/0040-5809\(75\)90020-9](https://doi.org/10.1016/0040-5809(75)90020-9)
- Watterson, J., & Alexander, J. (2004). *Southern flounder escapement in North Carolina* (pp. 1–3). Division of Marine Fisheries, Segments.
- Wickham, H. (2009). *ggplot2*. Springer.
- Willis, S. C., Hollenbeck, C. M., Puritz, J. B., Gold, J. R., & Portnoy, D. S. (2017). Haplotyping RAD loci: An efficient method to filter paralogs and account for physical linkage. *Molecular Ecology Resources*, 17(5), 955–965. <https://doi.org/10.1111/1755-0998.12647>
- Yang, Z., Chen, X., Zhao, N. A., Tang, H., Tao, J., Zhang, P., Shi, F., & Wan, C. (2018). The effect of different habitat types and ontogenetic stages on the diet shift of a critically endangered fish species, *Coreius Guichenoti* (Sauvage and Dabry de Thiersant, 1874). *International Journal of Environmental Research and Public Health*, 15(10), 2240. <https://doi.org/10.3390/ijerph15102240>
- Yeaman, S. (2015). Local adaptation by alleles of small effect. *The American Naturalist*, 186(S1), S74–S89. <https://doi.org/10.1086/682405>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: O'Leary, S. J., Hollenbeck, C. M., Vega, R. R., & Portnoy, D. S. (2021). Disentangling complex genomic signals to understand population structure of an exploited, estuarine-dependent flatfish. *Ecology and Evolution*, 11, 13415–13429. <https://doi.org/10.1002/ece3.8064>