Mixing rates in weakly differentiated stocks of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico

John S. Hargrove, Debra J. Murie, Daryl C. Parkyn, Emily V. Saarinen, James D. Austin

SEDAR70-RD-01

January 2020

Genetica (2018) 146:393–402
https://doi.org/10.1007/s10709-018-0031-1

ORIGINAL PAPER

Mixing rates in weakly differentiated stocks of greater amberjack (*Seriola dumerili*) in the Gulf of Mexico

John S. Hargrove[1,2] · Debra J. Murie[3] · Daryl C. Parkyn[3] · Emily V. Saarinen[1,4] · James D. Austin[1,2,3]

Received: 4 August 2017 / Accepted: 23 July 2018 / Published online: 25 July 2018

Abstract

The greater amberjack (*Seriola dumerili*) is a commercially and recreationally important marine fish species in the southeastern United States, where it has been historically managed as two non-mixing stocks (Gulf of Mexico and Atlantic). Mark-recapture studies and analysis of mitochondrial DNA have suggested the two stocks are demographically independent; however, little is currently known about when and where spawning occurs in Gulf of Mexico amberjack, and whether stock mixture occurs on breeding grounds. The primary objective of this study was to quantify stock mixture among breeding populations of amberjack collected from the Atlantic and Gulf of Mexico. Genetic data based on 11 loci identified very low, though statistically significant differentiation among Gulf of Mexico samples ($G_{ST} = 0.007$, $G'_{ST} = 0.009$; all $P = 0.001$) and between reproductive adults collected from two spawning areas ($G_{ST} = 0.007$, $G'_{ST} = 0.014$; all $P = 0.001$). Naïve Bayesian mixture analysis supported a single genetic cluster [$p(S|data) = 0.734$] whereas trained clustering (using Atlantic and Gulf spawning fish) gave the highest support to a two-cluster model ($p(S|data) = 1.0$). Our results support the argument that the genetic structuring of greater amberjack is more complex than the previously assumed two, non-mixing stock model. Although our data provide evidence of limited population structure, we argue in favour of non-panmixia among reproductive fish collected from the Gulf of Mexico and Florida Keys.

Keywords Admixture · Greater amberjack · Gulf of Mexico · Microsatellite · Stock mixing · Carangidae · Western Atlantic Ocean

Introduction

Connectivity and gene flow in marine organisms are fundamental evolutionary mechanisms that determine the distribution of genetic diversity among populations. Marine fishes have historically been thought of as large panmictic populations exhibiting minimal differentiation owing to the perceived lack of geographical barriers, high dispersal capability, and large effective population sizes (Hauser and Carvalho 2008). In recent decades, however, considerable advances have been made towards understanding how factors such as environmental features (e.g., Selkoe et al. 2008) and life history characteristics (e.g., Riginos and Liggins 2013) interact to structure genetic diversity. Evidence of complex genetic structure has been observed among an array of marine fishes (Young et al. 2015) and invertebrates (Truelove et al. 2015), and a cornerstone of successful stock management is an accurate understanding of population boundaries. Genetic data have been particularly useful in refining stock delineation and population assignments (Reiss et al. 2009), given their power to differentiate between historical and contemporary patterns of gene flow (Hellberg 2009). In fish species that experience populations mixing at various life history stages, utilizing genetic data to assess mixing rates between putative populations represents

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10709-018-0031-1) contains supplementary material, which is available to authorized users.

✉ James D. Austin
austinj@ufl.edu

[1] Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL, USA

[2] Present Address: Tennessee Cooperative Fishery Research Unit, Tennessee Tech University, Cookeville, TN, USA

[3] Program in Fisheries and Aquatic Sciences, School of Forest Resources and Conservation, University of Florida, Gainesville, FL, USA

[4] Present Address: Division of Natural Sciences, New College of Florida, Sarasota, FL, USA

a critical source of information that may serve to enhance management efforts and bolster our understanding of patterns of gene flow in the marine environment (Waples et al. 2008; Reiss et al. 2009).

The greater amberjack, *Seriola dumerili* (Risso 1810), is a large reef-associated carangid fish with a circumglobal, subtropical-temperate distribution. Greater amberjack is of commercial and recreational importance throughout their range and is a species of management concern along the Atlantic and Gulf coasts of the United States (Harris et al. 2007). Currently, greater amberjack is managed as two discrete stocks; fish in the Gulf of Mexico north of the Florida Keys (hereafter: Gulf) are managed under the Gulf of Mexico Fishery Management Council and fish along the Atlantic coast from Cape Hatteras, North Carolina, south to the Florida Keys (hereafter: Atlantic) are managed under the South Atlantic Fishery Management Council. The Gulf stock has been previously classified as overfished and undergoing overfishing, though this was not the case with the Atlantic stock (SEDAR 2008, 2014). Given that overfishing remains a concern, the determination of whether the Gulf of Mexico and Atlantic stocks are reproductively isolated is of utmost management importance.

Initial delineation of greater amberjack management stocks (Gulf versus Atlantic) was based on tag-recovery studies that revealed limited movement of tagged individuals between the two areas (McClellan and Cummings 1997). The genetic structure of greater amberjack is presently limited to a single study that inferred population structure using mitochondrial DNA (mtDNA) restriction fragment data (Gold and Richardson 1998). Patterns of spatial homogeneity and genetic differentiation (F_{ST}) between populations were interpreted as evidence of two genetic subpopulations, corresponding to the Gulf and Atlantic stocks. Despite molecular and tagging data that suggest a demographically independent population of greater amberjack occurring in the Gulf of Mexico, little is known about the spatial distribution of spawning efforts, or whether stock mixing (between Atlantic and Gulf) occurs on breeding grounds.

Greater amberjack from the Atlantic stock are thought to use a single spawning area off of South Florida and the Florida Keys, with peak spawning occurring in April and May (Harris et al. 2007). Recently, however, spawning aggregations of greater amberjack were located and sampled in the Gulf of Mexico off the coast of Louisiana (Murie et al. 2013; Smith et al. 2014). Reproductively active adults captured in Louisiana and tagged with pop-off satellite tags exhibited little net movement during the spawning season and, specifically, did not move to the Florida Keys spawning grounds (Murie et al. 2013). Therefore, there are at least two known spawning areas for greater amberjack, one off the coast of Louisiana in the Gulf of Mexico and another in the Atlantic Ocean off the Florida Keys. Regardless of the number of spawning areas available to Gulf amberjack, or their fidelity to specific sites, important questions are whether two biologically independent stocks are represented in the Gulf fishery, and if mixing occurs, what is the extent of gene flow among populations.

The explicit focus of this study was to test for evidence of genetic structuring among Gulf and Atlantic stocks of greater amberjack and to quantify the extent of mixing between populations in the Gulf of Mexico using nuclear genetic data. We used clustering algorithms to assign individual genotypes collected from fish at large in the Gulf of Mexico to reproductively active individuals collected from spawning grounds in Louisiana and the Florida Keys (i.e. reference samples) representing Gulf and Atlantic stocks, respectively. Our a priori hypothesis was that mixing between the Gulf and Atlantic stocks would be minimal based on historical tagging data, which suggests overall small-scale movement rates for individual greater amberjack.

Materials and methods

Sampling, DNA isolation, and genotyping

Greater amberjack were captured using hook and line and tagged as part of a movement study (Murie et al. 2013). A subsample of fish were examined for sex and reproductive condition based on oocyte staging as an indication of maturity and spawning stage (Smith et al. 2014). A 1 cm^2 portion of pectoral fin was removed from each fish and stored in 95% ethanol. A total of 543 greater amberjack were sampled from three regions in the Gulf of Mexico: (1) Louisiana (LA-GULF); (2) the Florida panhandle south to Apalachicola, FL (NE-GULF); and (3) the west coast of Florida (WFL-GULF); as well as from the Florida Keys (FK-ATL; Fig. 1). Sample collection occurred between February and June of 2008, with the exception of NE-GULF individuals that were sampled in March 2009. Fish sampled from the LA-GULF were divided into two sub-groups based on their reproductive state and were either non-reproductive if collected outside of the spawning season or diagnosed as sexually immature based on macroscopic inspection of gonads (designated as LA-GULF) or reproductive if fish showed evidence of spawning at the time of collection (LA-GULF-R; Murie et al. 2013). All fish collected from the Florida Keys were collected at known spawning grounds for the SE Atlantic greater amberjack stock (Harris et al. 2007) and were confirmed to be spawning at the time of collection (designated FK-ATL-R). The remaining sampling locations (NE-GULF and WFL-GULF) were not known to be specific spawning regions for greater amberjack.

DNA was extracted using the Gentra Puregene® Tissue Kit (Qiagen, Valencia, CA) following manufacturer's

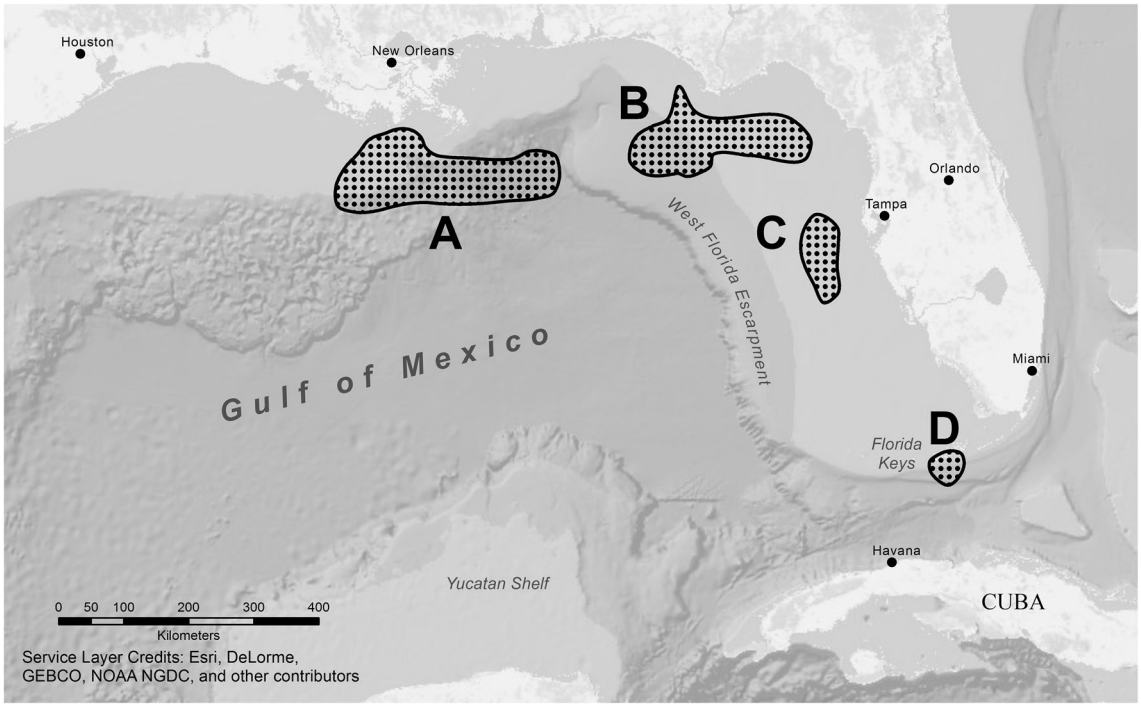

Fig. 1 Sampling regions for greater amberjack in the Gulf of Mexico, including: A LA-GULF (Louisiana), B NE-GULF (northeastern Gulf), C WFL-GULF (west-central coast of Florida), and D in the Atlantic Ocean off the Florida Keys (FK-ATL)

protocol and quantified using a Nanodrop spectrophotometer (Thermo Scientific). We generated genotypes using 15 unlinked microsatellite loci specifically designed for greater amberjack (Renshaw et al. 2006, 2007). PCR reactions were carried out in 15 µl simplex reactions containing 5.9 µl H_2O, 7.5 µl Qiagen multiplex PCR Mastermix (Qiagen, Valencia, CA), 0.1 µM M-13 labelled forward primer, 10 µM reverse primer, 10 µM M-13 dye-labelled primer (hexachlorofluorescein or 6-carboxyfluorescein), and 20 ng template DNA. All reactions were performed using thermocycling conditions of: 95 °C for 15 min; 35 cycles of 94 °C for 0.5 min, 58 °C for 1.5 min, 72 °C for 1.5 min; and a final extension at 72 °C for 10 min. Products were multi-pooled and run on an ABI 3130xl (Applied Biosystems, Foster City, CA) with a ROX 500 size standard (Applied Biosystems, Foster City, CA). Allele scoring was performed using GENEMARKER® software (SoftGenetics, State College, PA) and all allele calls were manually confirmed. We quantified genotyping and scoring error by re-genotyping 190 individuals for all loci.

Genetic analysis

Data were examined manually for outlier alleles, and reliability of genotype scoring and null alleles were evaluated using Micro-Checker version 2.2.1 (van Oosterhout et al. 2004). Deviations from Hardy–Weinberg equilibrium (HWE) and differentiation were evaluated using GenoDive version 2.0 (Meirmans and Van Tienderen 2004) and significance was determined following sequential Bonferroni correction of P values (Rice 1989). Tests for null alleles and HWE deviation were made among individuals sampled from the same region as this approach prevented spurious results due to population substructure (i.e. Wahlund effect). Global genetic differentiation among regional samples was estimated by computing the fixation index G_{ST} (Nei 1987), the standardized fixation index (G'_{ST}; Hedrick 2005) which controls for downward bias of G_{ST} in highly variable markers like microsatellites, and Jost's differentiation index (D; Jost 2008) that is independent of the amount of within-population diversity. Standard errors for differentiation estimates were obtained from a jackknife procedure over loci. We also conducted a permutation test (n = 1000) to determine if deviations were greater than expected under a random mating (panmixia) scenario. For comparative purposes, we also calculated pairwise F_{ST} among regional samples, with significance determined from 999 permutations. Finally, to evaluate differences in genetic structure among sexes we pooled sexes into two groups and tested for differences between male and female spawners ($F_{ST} > 0$) using 1000 bootstrap replicates.

Levels of heterogeneity between stocks may reflect the limited power of highly variable markers to detect genetic heterogeneity due to drift alone, versus under-sampling of allele frequencies. We used the program POWSIM 4.1

(Ryman and Palm 2006) to assess the statistical power of our analysis to reject the null hypothesis of genetic homogeneity between the two spawning populations. POWSIM simulates sampling from stocks under expected divergences assuming our two breeding aggregations (LA-GULF-R and ATL-R) diverged Wright–Fisher model without migration or mutation. An infinitely large initial population segregating for 11 loci with allele frequencies defined by our data was divided into 2 subpopulations (N = 90 and N = 40) of equal effective size (*Ne*) through random sampling of $2N_e$ genes. Each of the subpopulations of size N_e is allowed to drift for t generations, and the expected degree of divergence in generation t is then $F_{ST} = 1-(1–1/2N_e)^t$. Genetic homogeneity is tested using Fisher's exact test. Estimates of power were obtained as the proportion of significant outcomes when repeating the simulations 1000 times for each level of F_{ST} using default iteration parameters. We evaluated divergence times of 10, 25, 50, 80, and 100 generations and various effective populations sizes (1000, 4000 and 10,000). We also did an initial null simulation setting $t = 0$, to test that the initial number of false significances (α) was close to 0.05.

To evaluate the degree of mixing between stocks we utilized a stochastic optimization algorithm implemented in BAPS version 6.0 (Corander et al. 2004) which places individual genotypes into groups that correspond to latent genetic clusters. This approach can detect and distinguish between mixed or admixed groups of individuals (Corander et al. 2006) and has been shown to outperform other clustering algorithms when differentiation among samples is low (Latch et al. 2006) as was the case here. Briefly, BAPS generates a posterior probability for a number of genetic partitions given the data, $p(S|\text{data})$, based on a priori uncertainty of genetic mixture using a uniform prior that is restricted by an upper limit, K_{max} (an integer specified by the user), which represents the number of genetically panmictic partitions. For each replicate of a given K_{max} value, the optimal partition of individual genotypes into $\leq K_{max}$ is estimated, stored, and then merged according to log-likelihood scores once all K values have been explored.

First, we examined a range of maximum possible number of genetically divergent groups to examine whether increasing K_{max} affected posterior probabilities. We tested K_{max} at 5, 7, and 9. For each K we ran 10 replicates to allow the program to find the optimal cluster partitions within each K, as recommended by Corander et al. (2006). Second, we applied a 'trained clustering' (Corander et al. 2006) approach that incorporates a priori baseline data under the hypothesis that spawning aggregations in the north Gulf and Florida Keys represent at least quasi-independent genetic populations. This approach, when biologically justified, is useful for enhancing the statistical power of identifying the correct origin of individuals, particularly when genetic differentiation between populations is weak (Corander et al. 2006, 2008). This 'trained clustering' approach used LA-GULF-R and FK-ATL-R as training samples to guide assignment of the remaining Gulf samples. As with prior analyses, we tested K_{max} at 5, 7, and 9, using multiple replicates for each. Following mixture analysis, we post-processed BAPS output to evaluate the rate of admixture among genetic partitions, retaining genotypes having $P \leq 0.05$ of being admixed (Corander and Marttinen 2006). For each admixture analysis we ran 500 iterations per sample, using 50 reference individuals per population.

Next, we evaluated three alternative models of genetic structure in a Bayesian framework to describe the relationship among spawning individuals only (i.e. individuals from LA-GULF-R and FK-ATL-R that were characterized as reproductively active). The first model compared whether each spawning group consisted of a single genetic cluster or represented two distinct clusters. Each scenario in this model was given a uniform prior (0.5). The second model compared three competing scenarios: (1) the two groups consisted of two distinct genetic clusters, (2) all spawning fish belong to a single cluster, and (3) two genetic clusters existed but these groups were mixed based on empirically derived individual mixture results as described above. Each model was given a uniform prior probability (~0.333). Note that scenarios 1 and 2 of the second model reflect the same comparison as the first model. The third model was identical to the second, with the one exception being that priors were adjusted based on empirically derived mixing profiles (0.25, 0.25, and 0.50).

We utilized the software program ONCOR (http://www.montana.edu/kalinowski/) to estimate mixture proportions among sampled populations using a conditional maximum likelihood approach (Millar 1987). Mixture analysis generates a pairwise matrix of 'stock composition' estimates based on samples from a potentially mixed stock fishery, with outputs representing the proportion of a given stock (or in this case population sample) that were derived from a different stock. The accuracy of mixture analysis was evaluated via three-way error decomposition using the methods of Anderson et al. (2008). This analysis determines the percentage of the total error in assignment tests attributed to fishery sampling, genotypic sampling, and baseline sampling. Fishery sampling is error introduced by sampling too few fish from a fishery, genotyping error is due to sampling too few loci, and baseline sampling error is related to not knowing the true allele frequencies in a fishery. Error decomposition requires estimates of stock proportions (LA-GULF-R and FK-ATL-R) in each sample and a range of proportions (0.1–0.9) were evaluated to determine the impact on error decomposition.

We also performed assignment tests using ONCOR to estimate the population of origin for individual fish. The 'leave-one-out' test was used to evaluate the accuracy of

assignments to the breeding population of origin. Using only fish from LA-GULF-R and FK-ATL-R, this test sequentially removed each fish from the baseline and its origin was estimated from the remainder of the baseline stock. Each fish genotype was tested in this manner.

Results

Of the 543 greater amberjack sampled, we identified 40 reproductive fish from Louisiana (LA-GULF-R) and 91 individuals from the Florida Keys (FK-ATL-R). Sample sizes were greatest for the WFL-GULF ($N = 167$), followed by LA-GULF ($N = 136$) and NE-GULF ($N = 109$), none of which showed evidence of being reproductively active (i.e., no hydrated oocytes, no post-ovulatory follicles).

Genotype profiles were generated for all 543 greater amberjack. Significant deviations from HWE were observed for 26 of 75 total tests. Many of these deviations (n = 18) were specific to four loci (Sdu 32, 37, 46, and 16) with the remaining deviations (8) being distributed randomly across eight of the remaining 11 loci (i.e., in a non-systematic pattern). MICROCHECKER revealed potential HWE deviations due to null alleles in either 3 or 4 populations for markers Sdu 32, 37, 46, and 16. Null alleles were detected in at least one population (maximum 2) at 7 of the 11 remaining loci that deviated from HWE (all but Sdu 3, 5, 12). No tests inferred large allele dropout nor scoring error due to stuttering; however, based on comparisons among repeated amplifications, allelic dropout was high (range 4–10%) among the four loci that displayed evidence of both null alleles and HWE deviations (Sdu 32, 37, 46, and 16). As a result, we dropped these four loci for all analyses other than summary statistics. The average number of alleles per locus ranged from 15.13 to 18.67, while the effective number of alleles was considerably smaller (range 6.11–6.39). Observed levels of heterozygosity ranged from 0.689 to 0.729 (Table 1) and estimates of the inbreeding coefficient was highest for reproductive amberjack sampled from Louisiana ($G_{is} = 0.115$) and lowest for those sampled in the NE-GULF ($G_{is} = 0.021$). Genotypes are available from the corresponding author.

Overall ($G_{ST} = 0.007$, S.E. = 0.003) differentiation among samples in the Gulf of Mexico was low even after correction for biases associated with highly variable markers ($G'_{ST} = 0.009$, S.E. = 0.004; $D = 0.026$, S.E. = 0.011). All the observed values of genetic differentiation were higher than permutated data sets (all $P = 0.001$) which suggested low, yet statistically significant differentiation. Comparison between the two reproductive populations (LA-GULF-R versus FK-ATL-R) was similarly low ($G_{ST} = 0.007$, S.E. = 0.005; $G'_{ST} = 0.014$, S.E. = 0.010; $D = 0.047$, S.E. = 0.033) and all empirical values were highly significant (all $P = 0.001$). Outputs from tests of genetic differentiation suggested that greater amberjack from the Gulf of Mexico and Florida Keys do not represent a single, panmictic population.

Pair-wise values of F_{ST} also reflected significant allelic frequency differences among sample areas (Table 2). However, differentiation between males and females from spawning areas was not significantly different from zero ($F_{ST} = 0.006$, 99% CI −0.003 to 0.012).

POWSIM suggested that across a range of effective population sizes (N_e 1000–10,000) and number of generations

Table 1 Measures of population diversity for greater amberjack (*Seriola dumerili*) collected from the Gulf Mexico and Florida Keys, United States, based on 15 microsatellite loci

	N	n	n_e	H_o	H_e	G_{is}
LA-GULF-R (spawners)	40	15.13	6.13	0.689	0.779	0.115
LA-GULF (non-reproductive)	136	17.53	6.39	0.719	0.771	0.067
NE-GULF	109	16.20	6.11	0.729	0.745	0.021
WFL-GULF	167	18.67	6.35	0.674	0.748	0.099
FK-ATL-R (spawners)	91	15.93	6.14	0.688	0.761	0.096

Number of samples (N), number of alleles (n), and effective number of alleles (n_e) controlling for evenness of allele frequencies. Observed heterozygosity (H_o), expected (H_e) and inbreeding coefficient (G_{IS}) per sample population

Table 2 Pairwise differentiation values among amberjack sample locations

	LA-GULF-R	LA-GULF	NE-GULF	WFL-GULF	FK-ATL-R
LA-GULF-R (spawners)	–	0.001	0.001	0.001	0.001
LA-GULF (non-reproductive)	0.010	–	0.001	0.001	0.036
NE-GULF	0.015	0.007	–	0.001	0.001
WFL-GULF	0.016	0.007	0.014	–	0.001
FK-ATL-R (spawners)	0.011	0.002	0.008	0.003	–

F_{ST} values are on the lower matrix, P values are on upper matrix

(20–100), our loci had > 90% power to reject genetic homogeneity when F_{ST} ~ 0.011, the empirical pairwise differentiation between the LA-GULF-R and FK-ATL-R samples (Table 2). At F_{ST} ~ 0.005, we saw > 74% power, and power dropped below 70% only at 10 generations of drift at N_e = 4000, and at ~ 60 generations at N_e 10,000 (Supplemental File 1). Thus, our sample sizes and microsatellite markers should be adequate for detecting low levels of differentiation (e.g. $F_{ST} = 0.011$, Table 2) with power > 0.8 across most generation times and N_e values examined (Supplementary Material). Most F_{ST} values across simulated parameters were at or below observed, suggesting that sampling error is insufficient in explaining observed levels of differentiation.

For the naïve mixture analysis implemented in BAPS, the correct number of clusters (K) needed to describe the data was divided between $K = 5$ ($p(S|\text{data}) = 0.7638$) and $K = 8$ ($p(S|\text{data}) = 0.2362$). Of the ten best partitions visited (of 30 total examined), nine scenarios selected K = 5, and one was K = 8. Most genotypes (N = 540 of 543) were assigned to a singular partition, and two additional clusters consisted of 1 and 2 fish, respectively. The three identified clusters displayed little discernible structure; all three large partitions were distributed across the five sampling regions. Admixture was only detected with a single individual fish from the three main clusters (results not shown).

For the trained clustering mixture analysis, the optimal partition had fish distributed among 2 clusters ($p(S|\text{data}) = 1.0$). Estimated as groups of samples, these clusters corresponded to the NE-GULF clustering with LA-GULF-R, whereas WFL-GULF and LA-GULF clustered with FK-ATL-R. Individual-based mixture, however, identified less discrete groups (i.e., greater mixture; Fig. 2), and strong posterior support was observed for the presence of a single genetic cluster relative to other models relating the LA-GULF-R and FK-ATL-R populations (Table 3). These included comparisons of empirical results partitioning individual genotypes among clusters, regardless of prior (uniform or weighted toward empirical).

Attempts to identify the population of origin among reproductively active greater amberjack (i.e. 'leave-one-out' assignment test) returned 62.5% correct assignment for spawning fish sampled in Louisiana, and 77.8% correct assignment for those collected from the Florida Keys. Estimates of mixing among all greater amberjack sampling locations suggested high levels of mixing between populations in the Gulf of Mexico (LA-GULF, NE-GULF, and WFL-GULF) and the Florida Keys (range 0.706–0.809; Table 4). Mixing rates between reproductively active amberjack sampled in Louisiana (LA-GULF-R) and Gulf samples were on average much lower (range 0.191–0.294). Conditional likelihood mixture estimates of each area identified FK-ATL-R as the primary stock in each sample (Table 4). The error

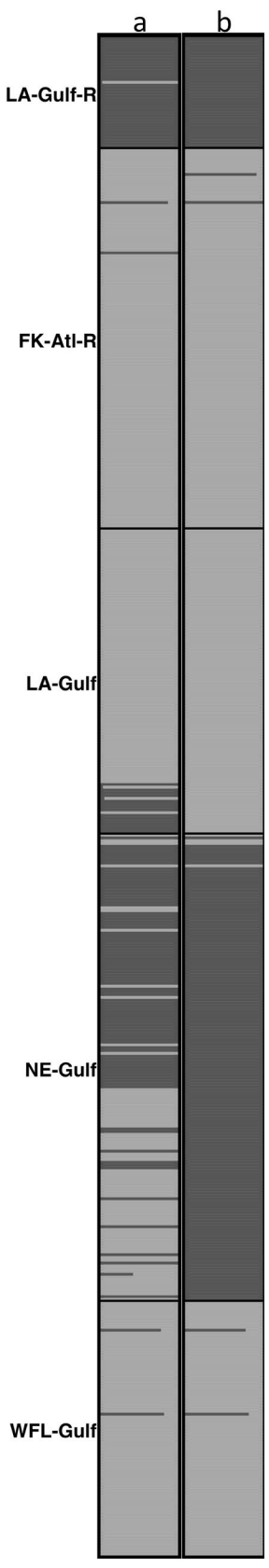

Fig. 2 Trained cluster analyses in BAPS of **a** individuals and **b** regional groups of greater amberjack sampled from the Gulf of Mexico and Florida Keys

Table 3 Bayes factor comparisons of models comparing alternative pre-specified clustering of reproductively active amberjack collected off the coast of Louisiana and the Florida Keys, Florida (LA-GULF-R and FK-ATL-R samples, respectively)

Model	Number of clusters	Prior	LA-GULF-R	FK-ATL	−LnL	Probability
1	Single	0.5	1	1	−5530.39	1
	Two	0.5	1	2	−5720.07	0
2	Single	0.33	1	1	−5530.39	1
	Two	0.33	1	2	−5720.07	0
	Empirical (two)	0.33	1 or 2	1 or 2	−5616.89	0

For model two, the results were identical when prior probabilities were weighted toward empirical (i.e. 0.25, 0.25, 0.5)

Table 4 Pairwise estimates of mixing for Gulf of Mexico greater amberjack based on genotypes sampled from baseline stocks [spawning fish caught off Louisiana (LA-GULF-R) and the Florida Keys (FK-ATL-R)] relative to non-reproductive fish sampled in the northern Gulf (LA-GULF), northeastern Gulf (NE-GULF) and the west coast of Florida (WFL-GULF)

Baseline stock	Area sampled			
	LA-GULF	NE-GULF	WFL-GULF	Combined
LA-GULF-R	0.191	0.294	0.173	0.211
FK-ATL-R	0.809	0.706	0.827	0.789

Table 5 Error decomposition derived from microsatellite data generated for greater amberjack collected from throughout the Gulf of Mexico

	Proportion Atlantic stock (Florida Keys)				
Baseline stock	0.1	0.3	0.5	0.7	0.9
Fishery	5.8%	31.8%	52.2%	27.7%	6.8%
Genotypic	2.1%	7.0%	9.9%	5.1%	2.0%
Baseline	92.1%	61.2%	38.0%	67.2%	91.2%

Fishery proportions tested range from 0.1 Atlantic to 0.9 Atlantic (versus a Gulf breeding population). Estimates are based on 10,000 simulated genotypes and a random fishery sample of 543

decomposition suggests that in most scenarios, the greatest contribution to error in detecting mixing is due to baseline error (Table 5). This suggested that considerably larger sample sizes (more so than greater numbers of loci) would be required to reliably estimate mixing rates, assuming genetic stock structure exists.

Discussion

In this study, we quantified connectivity among populations of greater amberjack sampled from the Gulf of Mexico and the Florida Keys using admixture analysis performed in a Bayesian framework and identified high levels of mixture among populations. The designation of greater amberjack management stocks have been assumed to reflect, to a considerable degree, independent demographic populations, an assumption not well supported by our genetic data. Our results indicate low levels of genetic differentiation among populations putatively identified as separate stocks, and pairwise estimates of mixing among populations suggest high rates of genetic exchange. Although our results do not resolve the issue of greater amberjack stock delineation, they highlight the complex nature of genetic mixing among greater amberjack.

Historical knowledge on connectivity between greater amberjack populations has been inferred from conventional tagging studies (i.e. physical tags such as dart or T-bar). Demographic exchange rates between the Gulf and Atlantic (SE U.S.) greater amberjack stocks estimated via mark-capture data are low (1.3–1.6%; Cummings and McClellan 1996; McClellan and Cummings 1997; Murie et al. 2013), and most tagged fish (Gulf or Atlantic) moved less than 100 nautical miles from their point of initial capture (Burch 1979). Murie et al. (2013) observed a mean movement distance of 69.54 ± 188.96 km (median distance = 8.0 km); however, they did observe a maximum distance of 1501 km as measured from a straight line. Combined, tagging study results suggest most individuals exhibit site fidelity (i.e., nearly resident) while select individuals wander widely. The observation of limited large-scale movements by greater amberjack appears to corroborate our genetic results; modest gene flow (i.e., handfuls of individuals per generation) between populations is realistic, and these movements may explain the low levels of genetic differentiation detected among populations (Waples 1998). It is important to note, however, that conventional tagging studies may belie more complex movement dynamics; specifically, their focus may be age specific and fail to capture dispersal that occurs during non-adult life stages such as those resulting from the drifting of pelagic larvae (Selkoe et al. 2008).

Patterns of genetic differentiation among populations of greater amberjack sampled from the Gulf of Mexico and Florida Keys were consistently low (global $G_{ST} = 0.007$), which reflects similar patterns from previous genetic studies. Gold and Richardson (1998) failed to identify significant heterogeneity in mtDNA haplotype frequencies of greater amberjack sampled from throughout the Gulf, or when Gulf samples were pooled with either Florida Keys

or Atlantic samples. Weakly significant heterogeneity was observed however when Florida Keys and Atlantic samples were pooled and contrasted with those from the Gulf, which served as the basis for the argument that two subpopulations of greater amberjack exist (i.e., the Gulf as independent from the Atlantic which includes the Florida Keys). These conclusions contrast with those of our Bayesian clustering analysis that strongly reflect a high rate of current and/or recent historical genetic mixing between individuals from the Gulf of Mexico and the Florida Keys. Studies of greater amberjack from other regions of the globe also suggest the presence of high rates of connectivity and mixing among populations; greater amberjack within the Mediterranean Sea fail to exhibit clear patterns of spatial heterogeneity (Ŝegvić-Bubić et al. 2016). Interestingly, two clades of greater amberjack were detected within the Mediterranean, yet individuals from these distinct clades were not separated spatially suggesting stock mixture but not contemporary gene flow. This latter observation highlights the potentially complex nature of population structure and movement patterns in greater amberjack potentially reflecting ancient historical events (e.g., colonization of newly available habitats).

Results from Bayesian cluster analysis differed depending on whether a subset or all greater amberjack populations were included. Naïve clustering performed on all sampling sites identified admixture between two or more latent genetic structures mixing within the Gulf of Mexico. In contrast, when the analysis was restricted to only spawning groups (i.e., only fish from Louisiana and Florida Keys that were reproductively active), results were uninformative in differentiating among two potential explanations. These two scenarios were either; (1) all individuals formed a single genetic cluster, or (2) that the two spawning groups were highly mixed assuming the two groups represent distinct genetic clusters. Our inability to differentiate between competing scenarios may in part be explained by the sensitivity of selected methods (clustering and detection of admixture) to handling moderate levels of gene flow (Latch et al. 2006). Prior simulation work has shown that in extreme scenarios of the isolation-connectivity continuum (i.e., zero migration or panmixia) there is commonly agreement between theoretical predictions and empirical observations, but in instances of weak or moderate differentiation these approaches perform less well (Latch et al. 2006; Waples and Gaggiotti 2006; Waples 2010). In spite of issues associated with model-based clustering of weakly differentiated populations (Putman and Carbone 2014), our results reflect high rates of current and/or recent historical genetic mixing across greater amberjack from the Gulf of Mexico.

The methods employed herein to evaluate genetic stock structure and mixing patterns, and previous genetic studies on greater amberjack (Gold and Richardson 1998), reveal limitations in applying neutral genetic methods to differentiate between Atlantic and Gulf of Mexico stocks of greater amberjack (see Putman and Carbone 2014 for review). Genotype approaches (assignment tests and clustering) can be more sensitive to fine-scale structuring than allele-based approaches (F_{ST}) (Garrick et al. 2010); however, these complementary approaches similarly failed to suggest strong differences between spawning fish collected from the putative Atlantic (Florida Keys) and Gulf of Mexico (Louisiana) stocks. Our coefficients of genetic differentiation (F_{ST} and analogues) reflect low levels of dissimilarity (i.e., ≤ 0.01) and suggest panmixia should be rejected. These metrics can be downwardly biased when applied to highly variable markers, something our unbiased estimators (G'_{ST}, D) helped to correct. Thus, the lack of genetic structure observed may reflect a true absence of differentiation over time scales of tens of generations (i.e., high migration and, by extension, high mixing rates; Waples 1998) or, alternatively, the effective population size (N_e) of greater amberjack could be considerably larger than our sampling regimen was able to differentiate (Hare et al. 2011). Both factors may be partially responsible for creating difficulty in evaluating stock structure in this and other mobile marine species (e.g. king mackerel, DeVries et al. 2002; Atlantic bluefin tuna; Taylor et al. 2011). Our power simulations suggest that, under reasonable estimates of N_e, that our markers probably perform well at detecting differentiation due to drift alone, and by extension, that our modest samples sizes probably are capturing accurate levels of differentiation. It is not known how long amberjack have been breeding in the northern Gulf of Mexico, and it may be that this is a relatively recent phenomenon, perhaps onset by increased artificial structure in that part of the Gulf. If true, and if the actual N_e of this breeding stock is still low, then our sampling is likely insufficient to accurately reflect differentiation, and rather is representing a sampling artefact. In such a case, the true F_{ST} would be considerably lower than 0.011 between breeding stocks (Table 2). That being said, increasing sample sizes for future microsatellite studies would be advantageous for evaluating sex-biased migration or stock mixing.

One assumption that remained untested until this study was the idea that the Florida Keys could be a breeding area for both Gulf and Atlantic stocks. The presence of large numbers of reproductively mature individuals in the northern Gulf of Mexico suggests greater amberjack utilize regions in the Gulf for spawning (Murie et al. 2013), an observation that raises the question of whether Gulf and Atlantic stocks are two demographically independent populations. Our results suggest the genetic structuring of greater amberjack is more complex than previously assumed and identify the possibility that high rates of historic gene flow may have occurred or that continued mixing among these two stocks may still be occurring. While our data do not include non-reproductive samples from the Atlantic stock,

it does argue against panmixia among reproductive fish in the Gulf and Florida Keys. An important consideration for future studies that seek to elucidate mixing rates and population structure among marine fishes in the southeastern United States is to sample beyond politically-defined boundaries (e.g., Jue et al. 2015). Furthermore, combining genetic and oceanographic models may help to elucidate movement patterns and stock structure among greater amberjack from Florida waters as well as those sampled across the Caribbean (e.g., Galindo et al. 2006). It is possible that the unique heterogeneity represented by the Florida Keys may reflect its importance as a mixed spawning area for multiple stocks of greater amberjack.

Acknowledgements This research was made possible by a Grant from the National Oceanic and Atmospheric Administration Fisheries Cooperative Research Program (#NA07NMF4540076). The study could not have been completed without the active collaboration of our associate investigators, including: Captain Myron Fischer (Cut-Off, LA; presently of the Louisiana Department of Wildlife and Fisheries), Captain Mark Hubbard (Madeira Beach, FL), Captain Ron Meyers (Little Torch Key, FL), and Captain Clay Bailey (Apalachicola, FL). We greatly appreciate them sharing their expertise in fishing for greater amberjack. This study required substantial field and laboratory support and we greatly appreciated the assistance of Geoff Smith, Doug Colle, Felipe Carvalho, and Chelsey Campbell throughout this study. Additional financial and logistically support was provided by the Program of Fisheries and Aquatic Sciences, School of Forest Resources and Conservation, Institute of Food and Agricultural Sciences at the University of Florida.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Anderson EC, Waples RS, Kalinowski ST (2008) An improved method for predicting the accuracy of genetic stock identification. Can J Fish Aquat Sci 65:1475–1486

Burch RK (1979) The greater amberjack, *Seriola dumerili*: its biology and fishery off southeastern Florida. MS thesis, University of Miami, Coral Gables, FL

Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. Mol Ecol 15:2833–2843

Corander J, Waldmann P, Marttinen P, Sillanpää MJ (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. Bioinformatics 20:2363–2369

Corander J, Marttinen P, Mäntyniemi S (2006) A Bayesian method for identification of stock mixtures from molecular marker data. Fish Bull 104:550–558

Corander J, Marttinen P, Siren J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinf 9:539–553

Cummings NJ, McClellan DB (1996) Fisheries information on the banded rudderfish, almaco jack, and lesser amberjack in the Gulf of Mexico through 1995. NMFS/SEFSC Miami Laboratory Contribution No. MIA-96/97-05. Available from National Marine Fisheries Service, Southeast Fisheries Science Centre, 75 Virginia Beach Drive, Miami, FL 33149, p 43

DeVries DA, Grimes CB, Prager MH (2002) Using otolith shape analysis to distinguish eastern Gulf of Mexico and Atlantic Ocean stocks of king mackerel. Fish Res 57:1–62

Galindo HM, Olson DB, Palumbi SR (2006) Seascape genetics: a coupled oceanographic-genetic model predicts population structure of Caribbean corals. Curr Biol 16:1622–1626

Garrick RC, Caccone A, Sunnucks P (2010) Inference of population history by coupling exploratory and model-driven phylogeographic analysis. Int J Mol Sci 11:1190–1227

Gold JR, Richardson LR (1998) Population structure in greater amberjack, *Seriola dumerili*, from the Gulf of Mexico and western Atlantic Ocean. Fish Bull 96:767–778

Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RSK, Ruegg K et al (2011) Understanding and estimating effective population size for practical application in marine species management. Conserv Biol 25:438–449

Harris PJ, Wyanski DM, White DB, Mikell PP, Eyo PB (2007) Age, growth, and reproduction of greater amberjack off the southeastern US Atlantic coast. Trans Am Fish Soc 136:1534–1545

Hauser L, Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish Fish 9:333–362

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59:1633–1638

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. Ann Rev Ecol Evol Syst 40:291–310

Jost L (2008) GST and its relatives do not measure differentiation. Mol Ecol 17:4015–4026

Jue NK, Brulé T, Coleman FC, Koenig CC (2015) From shelf to shelf: assessing historical and contemporary genetic differentiation and connectivity across the Gulf of Mexico in gag, *Mycteroperca microlepis*. PLoS One 10:E0120676

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE Jr (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conserv Genet 7:295–302

McClellan DB, Cummings NJ (1997) Preliminary analysis of tag and recapture data of the greater amberjack, *Seriola dumerili*, in the southeastern United States. In: Proc 49th Ann Gulf Carib Fish Inst, vol 49, pp 25–45

Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. Mol Ecol Notes 4:792–794

Millar RB (1987) Maximum likelihood estimation of mixed stock fishery composition. Can J Fish Aquat Sci 44:583–590

Murie DJ, Parkyn DC, Austin JD (2013) Seasonal movement and mixing rates of greater amberjack in the Gulf of Mexico and assessment of exchange with the South Atlantic spawning stock. SEDAR33-DW12. Southeast Data, Assessment and Review (SEDAR), North Charleston

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Putman AI, Carbone I (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecol Evol 4:4399–4428

Reiss H, Hoarau G, Dickey-Collas M, Wolff WJ (2009) Genetic population structure of marine fish: mismatch between biological and fisheries management units. Fish Fish 10:361–395

Renshaw MA, Patton JC, Rexroad CE, Gold JR (2006) PCR primers for trinucleotide and tetranucleotide microsatellites in greater amberjack, *Seriola dumerili*. Mol Ecol Notes 6:1162–1164

Renshaw MA, Patton JC, Rexroad CE, Gold JR (2007) Isolation and characterization of dinucleotide microsatellites in greater amberjack, *Seriola dumerili*. Conserv Genet 8:1009–1011

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

Riginos C, Liggins L (2013) Seascape genetics: populations, individuals, and genes marooned and adrift. Geogr Compass 7:197–216

Ryman N, Palm S (2006) POWSIM: a computer program for assessing statistical power when testing for genetic differentiation. Mol Ecol Notes 6:600–602

SEDAR (Southeast Data, Assessment, and Review) (2008) SEDAR 15: Stock assessment report 2 (SAR 2), South Atlantic greater amberjack. South Atlantic Fishery Management Council, North Charleston

SEDAR (Southeast Data, Assessment, and Review) (2014) SEDAR 33: Gulf of Mexico greater amberjack stock assessment report. South Atlantic Fishery Management Council, North Charleston

Šegvić-Bubić T, Marrone F, Grubišić L, Izquierdo-Gomez D, Katavić I, Arculeo M, Brutto SL (2016) Two seas, two lineages: how genetic diversity is structured in Atlantic and Mediterranean greater amberjack *Seriola dumerili* Risso, 1810 (Perciformes, Carangidae). Fish Res 179:271–279

Selkoe KA, Henzler CM, Gaines SD (2008) Seascape genetics and the spatial ecology of marine populations. Fish Fish 9:363–377

Smith GH, Murie DJ, Parkyn DC (2014) Nonlethal sex determination of the greater amberjack, with direct application to sex ratio analysis of the Gulf of Mexico stock. Mar Coast Fish 6:200–210

Taylor NG, McAllister MK, Lawson GL, Carruthers T, Block BA (2011) Atlantic Bluefin tuna: a novel multistock spatial model for assessing population biomass. PLoS One 6:e27693

Truelove NK, Ley-Cooper K, Segura-García I, Briones-Fourzán P, Lozano-Álvarez E, Phillips BF, Box SJ, Preziosi RF (2015) Genetic analysis reveals temporal population structure in Caribbean spiny lobster (*Panulirus argus*) within marine protected areas in Mexico. Fish Res 172:44–49

van Oosterhout C, Hutchison C, Shipley P, Wills D (2004) MicroChecker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J Hered 89:438–450

Waples RS (2010) Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. Mol Ecol Res 10:785–796

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419–1439

Waples RS, Punt AE, Cope JM (2008) Integrating genetic data into management of marine resources: how can we do it better? Fish Fish 9:423–449

Young EF, Belchier M, Hauser L, Horsburgh GJ, Meredith MP, Murphy EJ, Pascoal S, Rock J, Tysklind N, Carvalho GR (2015) Oceanography and life history predict contrasting genetic population structure in two Antarctic fish species. Evol Appl 8:486–509