Genetic population structure of hogfish (Labridae: *Lachnolaimus maximus*) in the southeastern United States

Seifu Seyoum*, Angela B. Collins* (contact author), Cecilia Puchulutegui*, Richard S. McBride[+], and Michael D. Tringali*

*Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, 100 Eighth Avenue SE, St. Petersburg, FL, 33701, USA

[+]NOAA Fisheries, Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA, 02543, USA

E-mail for contact author: angela.collins@MyFwc.com

Tele: 727-896-8626

Fax: 727-823-0166

Keywords: hogfish, genetic population structure, gene flow, microsatellite DNA.

ABSTRACT

Hogfish (Labridae: *Lachnolaimus maximus*) stock structure is poorly documented, confounding assessment of this reef fishery species. In this study, genetic structure was examined over a portion of the southeastern U.S. range of the species using a suite of 24 microsatellite DNA loci. Fin clips from 719 specimens were obtained from multiple geographic areas locations, ranging from Big Bend, Florida to the Carolinas. Genomic proportions of individual hogfish were partitioned into three distinct genetic clusters, geographically delineated as (1) the eastern Gulf of Mexico, (2) the Florida Keys and the southeast coast of Florida, and (3) the Carolinas. Clusters 1 and 2 converged along the coastal area west of the Florida Everglades, wherein the intervening samples comprised an admixture of individuals from adjacent clusters. Due to a discontinuity in sample coverage between cluster 2 and 3, the nature and precise location of the genetic break between the Florida Keys and the Carolinas requires further study. The presence of geographically-limited reproductive exchange in the hogfish leads us to recommend regionally partitioned collections and analyses of life-history and fishery data for stock assessments.

INTRODUCTION

A fundamental challenge to managing sustainable fisheries involves aligning biological evidence of stock structure with fishing and management sectors for the purpose of monitoring, assessment, and regulatory actions (Cadrin et al. 2014). This is particularly challenging around the Florida peninsula, where several zoogeographic regions overlap state (Florida) and federal boundaries, and fall under the jurisdiction of two federal fishery management councils (South Atlantic and Gulf of Mexico). In terms of zoogeography, there is a strong environmental gradient along (north-south) and between (east-west) coasts, resulting in distinctive faunal breaks at Cape Romano (west coast, Gulf of Mexico) and Cape Canaveral (east coast, Atlantic Ocean; Briggs and Bowen 2012). Subspecies are frequently recognized between the east and west coasts of Florida (Bowen and Avise 1990) and several coastal and marine species are considered to have distinct stocks on the respective coasts (Tringali and Bert 1996, Gold et al. 2002, McBride in press).

Many fishery species in the region remain data-poor with regard to stock structure. One such example is the hogfish (Labridae: *Lachnolaimus maximus*), a long-lived reef species that occurs in temperate to tropical waters of the western Atlantic Ocean and Gulf of Mexico (Claro et al. 1989, McBride and Richardson 2007, McBride et al. 2008). The species supports a modest commercial fishery in the southeastern United States, and is a highly valued target for recreational divers and spear fishers (McBride and Murphy 2003, FWC 2013, NMFS 2013). To date, no information on hogfish stock structure exists other than general behavioral and life history patterns and a preliminary genetic survey in the eastern Gulf of Mexico. Like most marine fishes, female hogfish broadcast eggs for external fertilization, which facilitates dispersal of the propagules away from the spawning site (Colin

1982). The planktonic larval duration lasts 3-5 weeks (Colin 1982, Victor 1986), which is fairly modest among reef fishes (Victor 1986, Soeparno et al. 2012), and does not imply extensive mixing of genotypes between ocean basins. Moreover, hogfish spawn in stable, site-specific harems throughout their range (i.e., they do not migrate or periodically aggregate to spawn) and there is specific evidence of reciprocal onshore larval dispersal and offshore movements by subadults (Muñoz et al. 2010, Collins and McBride 2011). These behaviors create demographic structure that may interact with stock structure.

Life history differences for hogfish (e.g., maximum age, maximum size and size and fecundity) exist between the eastern Gulf of Mexico and south Florida (McBride et al. 2008) and even within south Florida (McBride and Richardson 2007). These life history patterns were interpreted as a result of high fishing rates in more accessible habitat, and although they affect vital rates within each region, they have not been linked to underlying biological stock structure (McBride and Richardson 2007, McBride et al. 2008, McBride 2014). Questions remain, most recently as a request to review hogfish stock structure and unit stock definitions as part of the 2014 'Southeastern U.S. Hogfish Benchmark Assessment' (http://www.sefsc.noaa.gov/sedar/, SEDAR 37).

This paper introduces new genetic data to determine if hogfish stock structure exists within the coastal waters of the southeastern U.S., and if so, where the stock boundaries occur. Because genetic differentiation within marine populations can be difficult to detect, this investigation utilized microsatellite DNA markers, which possess a number of attributes that are well suited for such work relative to other marker types (e.g., allozymes and mitochondrial DNA) (Mariani and Bekkevold 2014, Antoniou and Magoulas 2014). Microsatellites are nuclear encoded, co-dominant markers that have characteristically high

mutation rates and, thus, a high degree of allelic variation. These loci are scattered throughout the genome and can be influenced independently by recombination, selection and drift; thus, each locus is expected to have its own slightly different genealogical history. Combining many loci makes a genomic sampling increasingly representative of the above genetic processes, and provides a more robust method for investigating gene flow and population connectivity (Hedrick 1999; Kalinowski 2002, 2005; Wilson and Rannala 2003). Here we apply microsatellite loci previously isolated for hogfish and optimized for routine assay (Seyoum et al. 2012) to survey specimens collected from Big Bend, Florida to the Carolinas.

MATERIALS AND METHODS

Sample collection and DNA extraction

Study specimens (n = 719) were collected through intercepting recreational and commercial spear fishers or during directed research trips by biologists from the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute (FWC FWRI). Fin clips were removed and preserved in 70% ethanol. Total DNA was isolated from approximately 500 mg of fin clip tissue using PUREGENE DNA isolation kits (Gentra Systems, INC., Minneapolis Minnesota) and rehydrated in 50 µl of deionized water.

Collection locations were subdivided into nine geographic areas, referred to hereafter as 'samples' (Figure 1). These study samples were established predominantly by latitude and coast (west [Gulf of Mexico], east [Atlantic Ocean]) to delineate areas corresponding to recognized faunal breaks, major estuaries, and, on a broader scale, management jurisdictions. For example, faunal breaks occur at Cape Romano between samples 5 and 6, Cape Sable

between samples 6 and 7, and Cape Canaveral between samples 8 and 9. Considerable estuarine flow onto the continental shelf occurs from the Suwannee River between samples 1 and 2, Tampa Bay (sample 3) and Charlotte Harbor (sample 4). In terms of management jurisdictions, the Gulf of Mexico Fishery Management Council regulates federal fisheries throughout the Gulf of Mexico (samples 1-6); the South Atlantic Fishery Management Council regulates fisheries in federal waters from the Carolinas to the Florida Keys (samples 7-9); and Florida regulates state fisheries inside 9 miles in the Gulf of Mexico and inside of 3 miles in the Atlantic (samples 1-8).

Microsatellite genotyping

Specimens were genotyped using 24 of 29 microsatellite markers identified in Seyoum et al. (2012); (Lmax11, Lmax14, Lmax15, Lmax24, and Lmax31 markers were not used). Multiplex PCR amplifications were carried out in an Eppendorf thermal cycler containing 50-100 ng of total DNA, 10 µl 50uM of dNTP mix, 0.25 µl of 0.1mg/ml BSA, a combination of three optimally selected primers of three loci with each forward primer labeled with different fluorescent dye, 5 µl of *Taq* Polymerase Buffer (10×) containing 15mM $MgCl_2$ (Promega) and 1.25 units of *Taq* polymerase (Promega). The reaction profile was 94°C for 2 min, 35 × (94°C for 35 s, 55°C for 35 s, 72°C for 35 s) and final extension at 72°C for 30 min. Fragments were visualized on an ABI 3130 XL genetic analyzer and genotyped using GeneMapper software (version 4.0; Applied Biosystems, Inc.). For fragment assays, we used Gene Scan-500 ROX-labeled size standard (Applied Biosystems, Inc. Foster City, CA).

Data analysis

GENEPOP data files were generated from fragment sizes recorded using the Microsatellite Marker Toolkit add-on (version 3.1.1; Park 2001): http://animalgenomics.ucd.ie/sdepark/ms-

toolkit/); these were converted to other formats as needed using the conversion tool PGDspider (version 2.0.1.9; Lischer and Excoffier 2012). Pairwise genetic distances (F_{ST}) between samples (Weir and Cockerham 1984) were estimated with 10,000 permutations in the program GENETIX (Belkhir et al. 2000). Departures from Hardy-Weinberg equilibrium (HWE) were determined by using GENEPOP (version 3.4; Raymond and Rousset 1995, Rousset 2008). Sequential Bonferroni corrections were applied to multiple tests of hypotheses (Rice 1989). Observed (H_o) and expected heterozygosity (H_e, with and without a bias correction), averaged over all loci, were obtained from GENETIX (Belkhir et al. 2000). Null allelism was investigated with the randomization test of Guo and Thompson (1992) and the U-test statistic of Raymond and Rousset (1995), using the program ML-NullFreq (http://www.montana.edu/kalinowski/software.htm). For each locus, microsatellite variation was quantified in terms of genetic diversity, number of alleles, and allelic richness (a diversity measure that corrects for differences in sample size) using the program FSTAT (version 2.9.3.2; Goudet 2001). Chi-square tests were performed to determine if samples differed significantly with respect to the above standard genetic measures.

Phenetic clustering

Corrected and uncorrected F_{ST} genetic distances (Latter 1972) were generated from allele frequencies and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) phenograms based on these distances were reconstructed using the Software POPTREE2 (Takezaki, Nei, and Tamura 2010) and visualized using the software FIGtree (http://tree.bio.ed.ac.uk/software/figtree/). Statistical significance of relationships was assessed via 1000 bootstrap replications. Adopting patterns observed in the phenogram, samples were categorized into geographic clusters for further analysis of genetic structure.

Genetic structure

Genetic structure among the study samples was examined using three analytical approaches. The first was based on the analysis of molecular variance (AMOVA) as implemented in the program ARLEQUIN (version 3.5.1.3; 100,000 permutations, Excoffier and Lischer 2010). AMOVA follows an *a priori* hierarchical approach in which correlations among genotypes at various levels are partitioned as F-statistics. The proportion of variation among groups (F_{CT}), within groups (F_{SC}), and within samples (F_{ST}) was computed and the φ statistic assessed by the permutation method of Excoffier et al. (1992). Initially, the *a priori* hierarchical structure analyzed was based on the results of above phenetic reconstruction. For comparison, sample groupings other than that indicated by the phenogram were also considered.

The second analytical approach was based on the Bayesian population-assignment algorithm as implemented in the program STRUCTURE (version 2.3.4; Pritchard et al. 2000). With this algorithm, individuals were probabilistically and proportionally assigned to one or more genetic clusters (*K*) in a manner that minimized Hardy-Weinberg and linkage disequilibrium among their multilocus genotypes. For $K = 1$ through 9, ten simulations were conducted using 2E6 Markov-Chain Monte Carlo (MCMC) replicates following a 1E6 burn-in period. We adopted the admixture model and the independent allele frequency option to minimize the chance of overestimating the number of clusters present in the data (Pritchard et al. 2009). We used STRUCTURE HARVESTER (version 0.6.93; Earl and vonHoldt 2012) with each of the above replicate runs to compute the *ad hoc* statistics *Ln* (*K*) and Δ*K* in order to determine the most plausible 'base' value of *K* clusters (i.e., the upper-level hierarchy). *Ln* (*K*) denotes the log probability of the data at a given modeled *K* value; Δ*K* is based on the rate of change in *Ln* (*K*) between successively modeled *K* values. Simulation studies (Evanno et al. 2005) have shown that Δ*K* provides the most accurate indication of genetic

structure under a variety of modeling conditions. We then used CLUMPP (version 1.1.2; Jakobsson and Rosenberg 2007) to determine the optimal alignment for replicate analyses and mean genomic membership coefficients across replicate runs for samples and individuals. The sample coefficients were plotted using EXCEL (Microsoft) and the 719 individuals using DISTRUCT (version 1.1; Rosenberg 2004) and visualized using the program Ghost View (http://pages.cs.wisc.edu/_ghost).

Finally, to determine if genetic relationships among the samples conformed to a pattern of isolation-by–distance gene flow (Malécot 1955; Wright 1943), we computed the mantel correlation coefficient R between genetic distance (F_{ST}) and geographic distance (Km) using the program GenAlEx (version 6.5; Peakall and Smouse 2006, 2012). Significance of the correlation coefficient was tested via 9000 random permutations.

RESULTS

Standard genetic measures and distances

Significant heterozygote deficiencies were sporadically detected at three loci (Lmax4, Lmax29 and Lmax 35) in one to five samples when biased estimates of expected heterozygosities were considered; however, no deviations were observed when unbiased estimates were considered. Presumptive frequencies of null alleles at those loci exhibiting heterozygote deficits ranged from 0.17 to 0.24. There were no significant differences among samples in the mean values of standard genetic measures (Table 1). For all loci, a total of 350 unique alleles were identified and of these, an average of 216 alleles occurred in each sample. Over the 36 possible pairwise comparisons, 30 sample pairs had genetic distances that were significantly greater than zero ($P<0.05$).

212

213 **Phenetic clustering**

214 UPGMA reconstructions based on F_{ST} distances (Figure 2) revealed three geographically
215 oriented but poorly supported clusters. Cluster 1 included specimens collected in the eastern
216 Gulf of Mexico from the Florida Panhandle (Big Bend) to Naples (samples 1 – 5). Cluster 2
217 included specimens collected from the Florida Keys and along the southeastern coast of
218 Florida (samples 7 – 8). Cluster 3 included specimens collected from the Carolinas (sample
219 9). Significant bootstrap support was observed only for the branch separating sample 9
220 (Cluster 3) from the remaining samples (clusters). The position of the sample 6 (Everglades
221 region) was unresolved in the reconstruction, fluctuating between Cluster 2 (Figure 2A;
222 uncorrected F_{ST}) and Cluster 1 (Figure 2B; corrected F_{ST}).

223

224 **Analysis of Molecular Variance**

225 In the AMOVA, 98.17% of the variation occurred within samples and 1.83% among
226 samples. The overall F_{ST} value of 1.83 (P<0.0001) was indicative of significant
227 differentiation among samples due to the presence of spatial structure at both regional and
228 local scales. The highest among-group variance in the AMOVA (Table 2) was observed
229 when samples were grouped according to the phenetic clusters shown in Figure 2B.
230 However, translocation of sample 6 from Cluster 1 to Cluster 2 only slightly reduced the
231 among-group variance.

232

233 Pairwise F_{ST} values between uncorrected (and corrected) F_{ST} phenetic clusters, respectively,
234 were 0.015 (0.016) between Clusters 1 and 2, 0.027 (0.023) between 1 and 3, and, 0.030
235 (0.039) between 2 and 3 (Table 2). All pairwise F_{ST} comparisons among the clusters were
236 highly significant (P<0.001). Overall, the F_{ST} comparisons and the AMOVA confirmed the

presence of three geographically based clusters, but these appeared to be hierarchically arranged with the least genetic differentiation between Clusters 1 and 2, and the differentiation between Clusters 2 and 3 exceeding that observed between the geographically disjunctive Clusters 1 and 3. Thus interestingly, the observed pattern of genetic differentiation did not conform to isolation-by-distance expectations over the studied range.

Bayesian population assignment

In the absence of a methodology to determine whether values of $Ln(K)$ statistically differ, inferences were derived herein by evaluating replicate likelihoods and resultant ΔK statistics for different values of K. $Ln(K)$ values increased quickly from $K = 1$ to $K = 2$, and somewhat less quickly from $K = 2$ to $K = 3$, before reaching a plateau at successive values (Figure 3A). The largest value of ΔK occurred for $K = 2$ ($\Delta K = 281$) and, secondarily, for $K = 3$ ($\Delta K = 118$) (Figure 3B). At the base hierarchical level of K=2, members of the phenetically delineated Clusters 1 and Cluster 3 were predominantly conjoined within a single Bayesian cluster whereas members of phenetic Cluster 2 were predominantly assigned to a discrete Bayesian cluster (Figure 4A). At the secondary hierarchical level of $K = 3$, members of phenetic Clusters 1 and 3 were assigned to separate Bayesian clusters (Figure 4B). The CLUMMP analysis indicated that samples 5 and 6, respectively, were admixtures of Bayesian Clusters 1 and 2, exhibiting graduated mean genomic proportions (Figure 4C).

Mantel Test

There was only a sufficient number of paired comparisons within Cluster 1 (Big Bend to Everglades) to perform a within-cluster Mantel test. There was no significant correlation between genetic and geographic distances ($p = 0.125$, $R = 0.334$) (Figure 5).

In summary, the phenetic clustering, AMOVA, and Bayesian population-assignment analyses elucidated a concordant pattern of genetic structure within the studied range. Three geographically based clusters were delineated as (1) the eastern Gulf of Mexico, (2) the Florida Keys and the southeastern coast of Florida, and (3) the Carolinas (Figure 6). There was no indication of isolation-by-distance gene flow within the eastern Gulf or over the sampled range. The area west of the Florida Everglades (sample 6) appeared to serve as a genetic break between the respective stocks. Due to a discontinuity in sample coverage, the nature and precise location of the apparent genetic break between the Florida Keys and the Carolinas requires further study.

DISCUSSION

At the outset of this study, three genetic stocks of hogfish could not be confidently predicted, but neither were they unexpected at this landscape scale. Capes Romano and Canaveral have previously been identified as faunal breaks, associated with major shifts in marine animal communities. Briggs and Bowen (2012) identify Cape Romano as separating the Gulf of Mexico from the Caribbean province, and Cape Canaveral as separating the Carolinian province from the Caribbean province. Strong genetic differences between the two Florida coasts are known for several inshore and nearshore fishes including common snook, (*Centropomus undecimalis*) (Tringali and Bert 1996), red drum (*Sciaenops ocellatus*) (Gold et al. 1999; Seyoum et al. 2000; Gold and Turner 2002), and spotted seatrout (Seyoum et al. 2014). In contrast, only weak genetic differences, if any, have been found between Florida coasts for the reef fishes such as red grouper (*Epinephelus morio*) and scamp (*Mycteroperca phenax*) (Zatcoff et al. 2004), vermilion snapper (Tringali and Higham 2007), and for the pelagic king mackerel (*Scomberomorus cavalla*) (Gold et al. 2002). Based upon existing

287 data, hogfish stock structure and levels of population connectivity in the southeastern United
288 States are generally more similar to inshore and nearshore fishes than to offshore reef fishes
289 or large coastal pelagics.
290
291 Planktonic larval duration is a good first approximation of the potential for gene flow in
292 marine systems (Waples 1987, Purcell et al. 2006, Soeparno et al. 2012). Hogfish eggs and
293 larvae are transported away from the reef habitats where spawning occurs, and juveniles
294 settle within shallow, nearshore grass beds (Davis, 1976; Colin 1982). They have a modest
295 planktonic larval duration lasting approximately 35 days before strong benthic orientation
296 begins (Colin 1982, Victor 1986). Post hatch larvae are found under floating material, and
297 may also form bubbles around themselves while floating near the surface (Colin 1982),
298 suggesting that distribution of larvae would be influenced more heavily by surface currents
299 than mid-water or benthic transport.
300
301 Nonetheless, the hydrodynamics around Florida are complex enough that predicting the
302 outcome of the average flow, in the average direction, for the average planktonic period is too
303 simplistic. Along the west Florida shelf, current flows are weakest near Cape Romano, which
304 may present a barrier to gene flow between clusters 1 and 2; however, current direction above
305 the shelf also shifts from southerly to northerly throughout the year, and at least near the shelf
306 break, a strong looping current interacts periodically with local wind forces above the shelf
307 (He and Weisberg 2003). The peak spawning period for hogfish in the eastern Gulf of
308 Mexico (March – May; McBride et al. 2008; Collins and McBride, in review) occurs at a
309 time when the currents shift from a predominant southeast flow to a northwestern flow (Liu
310 and Weisberg 2005, 2012). However, spawning occurs to some extent in almost all months of
311 the year (except September; Collins and McBride, 2011, in review). Spawning throughout the

312 year would subject larvae to a diverse (and difficult to predict) set of conditions, both
313 physical (e.g., seasonal shifts in currents, temperature) and biological (e.g., prey availability
314 through seasonal plankton blooms) that affect dispersal vectors and survival rates. The Loop
315 Current in the Gulf of Mexico is an upstream portion of the Gulf Stream and impacts portions
316 of the west Florida shelf to differing degrees throughout the year. The Loop Current
317 presumably occurs beyond the range of hogfish spawning activity on the west Florida shelf
318 (Cluster 1; spawning < 70 m; Collins and McBride, in review), but the effects of this major
319 current and its associated eddies are difficult to assess since many regional ichthyoplankton
320 surveys report taxa to only the family level (e.g., Labridae; Richards et al. 1993). Along the
321 Atlantic side of the Florida Keys, the dominant current flow is to the east, fed by the Florida
322 Current that flows through the Florida Strait and into the Gulf Stream. This current flow
323 suggests that the majority of larvae produced south of the Everglades region would be
324 transported into the Keys or Atlantic Ocean with very little larval transport north along the
325 west Florida shelf. Models of regional hydrodynamics combined with the increased
326 information available regarding hogfish distribution and spawning could simulate larval
327 dispersal and settlement around Florida. Elsewhere, biological data are lacking; for example,
328 courtship of hogfish has been observed off of North Carolina (Parker 2000), but the timing
329 and extent of spawning in this region (Cluster 3) has not been described.
330
331 Adult behavior may also play an important role in determining connectivity and population
332 structure (Frisk et al., in press) All hogfish mature as a female first, but if they live long
333 enough they will eventually change sex (McBride and Johnson 2007). A single male (which
334 may have been previously a female in the same harem) will spawn daily with multiple
335 females in harems containing as many as 15 individuals (Colin 1982, McBride and Johnson
336 2007, Muñoz et al. 2010; Collins and McBride 2011). Mature hogfish have relatively strong

site fidelity to the same reef habitats for months and even years (Colin 1982, Lindholm 2006, Muñoz 2010). However, fish less than 2 years old have rarely been observed from offshore habitats (Collins and McBride 2011) and because juveniles settle in shallow, inshore habitats, it is assumed that they gradually move offshore with growth. Although natal homing is not predicted for hogfish, it has been demonstrated for other reef fish (Paris et al. 2013); at the very least, such reciprocal ontogenetic movements demonstrated by hogfish focus gene flow in an inshore-offshore direction rather than along-shore direction.

Demographic differences in hogfish between or within regions of Florida have been noted (McBride and Richardson 2007, McBride et al. 2008, Collins and McBride 2011). In the central eastern Gulf of Mexico, adult size, longevity, and fecundity differ between deep and shallow water fish across the shelf (Collins and McBride 2011), unrelated to genetic structure (Seyoum et al. 2012). Although demographic differences also exist between the Gulf of Mexico and south Florida (McBride and Richardson 2007, McBride et al. 2008), which can now be said to be genetically different, it is premature to assume these life history differences are a result of genetic differences. The effect of fishing on size and at age, maturity, fecundity, and harem stability is very likely playing a large role in the demographics of hogfish throughout their range (McBride et al. 2008, Muñoz et al. 2010). In addition to fishing pressure, the potential effects of red tide on recruitment rates should be considered together with this new information on genetic stock structure around Florida (McBride and Richardson 2007, Collins and McBride 2011).

If the inferred close relationship between Clusters 1 and 3 is real, the genetic forces responsible are not readily apparent. It could be possible that hogfish in Clusters 1 and 3 are connected to some degree by gene flow to a larger Caribbean stock via the Gulf Stream

362 current (CITE), with those in Cluster 2 being spatially isolated due to localized
363 hydrodynamics and/or spawning behaviors. Alternatively, it is possible that for unknown
364 demographic reasons genetic drift has had a greater relative effect on the neutral genetic
365 composition of Cluster 2. Finally, it is possible that temperature, or some other correlate
366 along a latitudinal cline, could be driving selection for particular genotypes. Temperatures
367 are high (> 25 C) and relatively homogeneous around Florida during the summer (June-
368 September), but the temperatures over the western Florida shelf are much colder than in south
369 Florida (Cluster 2) during the rest of the year (He and Weisberg 2003). Such a latitudinal
370 cline could result in the closer genetic relationship between Clusters 1 and 3, where yet
371 unknown environmental effects may be selecting for a more similar genotype. Additional
372 study is needed to address these hypotheses.
373
374 Application of our findings to monitoring, assessment, and regulatory actions should not be
375 problematic. Landings and fishing effort in Florida are already collected at the county level
376 (FWC 2013), and the Florida Keys are contained with a single county (Monroe). Federal
377 fisheries data can also be separated by coast (e.g., NMFS 2013) and we predict that regional
378 assessments will further demonstrate differences in fishing pressures between regions: fishing
379 pressure within Cluster 1 is high (annual landings data reference for the Gulf, NMFS 2013)
380 but the pressure is more spread out due to the nature of the habitat and the expansive west
381 Florida shelf. Fishing pressure is also high in Cluster 2, where it is condensed into a much
382 smaller area that is more accessible to divers. In both regions, cluster 1 and 2, hogfish
383 residing in shallow depths are cropped out of the population as soon as they reach legal sizes,
384 particularly in the most (human) populated areas (McBride and Richardson 2007). Cluster 3
385 presumably has lower fishing pressure; there are commercial and recreational fisheries in this
386 area, but the travel time, distance from shore, weather patterns and lower human population

density likely reduce fishing pressure for this species off of the Carolinas. Regional assessments that can quantify fishing effort and landings have been challenging in the past (Kingsley 2004). As an alternative, Ault et al. (2005) promoted a size-based approach to addressing the data poor nature of the hogfish assessment, and Collins and McBride (2011) underscore considerations of spatial demographic structure. The genetic data presented herein strongly suggest that regional assessments are warranted for the eastern Gulf of Mexico, south Florida, and a broad but as yet undefined area from northeast Florida to the Carolinas.

Acknowledgements

We thank the staff of the Fisheries-Independent and Fisheries-Dependent Monitoring Programs of the Fish and Wildlife Research Institute for providing samples. We are especially grateful to Jim Atack, Erin Burge, Chip Collier, Kenny Fex, Jon Haag, Justin Shepard, Albie Solana, and Matthew Stokley for providing samples from North and South Carolina, Elizabeth Schotsman for providing so many individuals from the Florida Keys and to John Herrera for providing east Florida samples. We would like to thank Captains B. Bateman, L. Borden, W. Butts, J. DeLaCruz, C. Gardinal, C. Grauer, T. Grogan, B. Hardman, J. Hermes, S. Hooker, I. Lathrop, S. Lucas, K. Ludwig M. Joswig, D. O'Hern, D. Palmer, H. Scarboro, C. Schnur, R. Taylor and the contributing participants of the St. Pete Open, Spearboard Open, Southern Open, Key West Open and the Wrightsville Beach Spearfishing Tournaments. Kyle McWhorter and Jamie Williams assisted with data entry. The majority of the work described herein was funded by grant NA05NMF4540040 to the Florida Fish and Wildlife Conservation Commission from the National Oceanic and Atmospheric Administration's Cooperative Research Program, as well as through the U.S. Department of the Interior, U.S. Fish and Wildlife Service, under the Federal Aid in Sport Fish Restoration

Program, Grant F-69. The statements, findings, and conclusions are those of the authors and do not necessarily reflect the views or policies of the Department of Interior. Mention of trade names or commercial products does not constitute their endorsement by the U.S. government.

References

Antoniou, A., and A. Magoulas.

2014. Application of mitochondrial DNA in stock identification. Pages 257–285 *in* S. X. Cadrin, L. A. Kerr, and S. Mariani, editors. Stock Identification Methods (Second Edition). Academic Press, San Diego.

Ault, J. S., S. G. Smith, and J. A. Bohnsack.

2005. Evaluation of average length as an estimator of exploitation status for the Floridacoral-reef fish community. ICES J. Mar. Sci. 62:417–423.

Belkhir, K, P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme.

2000. GENETIX, Logiciel sous WindowsTM pour la génétique des populations. Laboratoire Génome, Populations, Interactions CNRS UMR 5000, Université de Montpellier II, Montpellier France.

Bowen, B. W., and J. C. Avise.

1990. Genetic structure of Atlantic and Gulf of Mexico populations of sea bass, menhaden, and sturgeon: influence of zoogeographic factors and life-history patterns. Mar. Biol. 107:371–381.

Briggs, J. C., and B. W. Bowen.

2012. A realignment of marine biogeographic provinces with particular reference to fish distributions. J. Biogeogr. 39(1):12–30.

Cadrin, S. X., L. A. Kerr, and S. Mariani.
2014. Stock Identification Methods (Second Edition). Academic Press, San Diego.
Claro, R., A. Garcia-Cagide, and R. Fernández de Alaiza.
1989. Caracteristicas biológicas del pez perro, *Lachnolaimus maximus* (Walbaum), en el golfo de Batabanó, Cuba. Rev. Inves. Mar. 10(3):239–252.
Colin, P. L.
1982. Spawning and larval development of the hogfish, *Lachnolaimus maximus* (Pisces: Labridae). Fish. Bull. U. S. 80(4):853–862.
Collins, A. B., and R. S. McBride.
2011. Demographics by depth: Spatially explicit life-history dynamics of a protogynous reef fish. Fish. Bull. U. S. 109:232–242.
Collins, A. B., and R. S. McBride.
in review. Spatial and seasonal variations of reproductive potential in a reef fish. Fish. Manage. Ecol..
Davis, J.C.
1976. Biology of the hogfish, *Lachnolaimus maximus* (Walbaum), in the Florida Keys. Master of Science Thesis. University of Miami, Coral Gables, FL 86 pp.
Earl, D. A., and B. M. vonHoldt.
2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Cons. Gene. Res. vol. 4 (2) pp.359–361. doi: 10.1007/s12686–011–9548–7.
Evanno, G., S. Regnaut, and J. Goudet.
2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14:2611–2620.
Excoffier, L., and H.E. L. Lischer.

2010. Arlequin suite (version 3.51.3) A new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10:564–567.

Excoffier, L., P. E. Smouse, and J. M. Quattro.

1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: applications to human mitochondrial DNA restriction data. Genetics 131:479–491.

Frisk, M. G., A. Jordaan, and T. J. Miller.

in press. Moving beyond the current paradigm in marine population connectivity: are adults the missing link? Fish Fish. http://dx.doi.org/10.1111/faf.12014.

Florida Fish and Wildlife Conservation Commission.

2013. Species account: Hogfish (*Lachnolaimus maximus*) in Florida. (http://myfwc.com/research/saltwater/status-trends/finfish/hogfish/).

Gold, J., and T. Turner.

2002. Population structure of red drum (*Sciaenops ocellatus*) in the northern Gulf of Mexico, as inferred from variation in nuclear-encoded microsatellites. Mar. Biol. 140(2):249–265.

Gold, J. R., E. Pak, and D. A. DeVries.

2002. Population structure of king mackerel (*Scomberomorus cavalla*) around peninsular Florida, as revealed by microsatellite DNA. Fish. Bull. U. S. 100:491–509.

Gold, J. R., L. R. Richardson, and T. F. Turner.

1999. Temporal stability and spatial divergence of mitochondrial DNA haplotype frequencies in red drum (*Sciaenops ocellatus*) from coastal regions of the western Atlantic Ocean and Gulf of Mexico. Mar. Biol. 133(4):593–602.

Goudet, J.

2001. FSTAT, A program to estimate and test gene diversities and fixation indices (version 2.9.3.2) Available from http://www2.unil.ch/popgen/softwares/FSTat.htm.

Guo, S.W. and E.A. Thompson.

1992. Performing exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics 48:361–372.

He, R. Y., and R. H. Weisberg.

2003. A loop current intrusion case study on the West Florida shelf. J. Phys. Oceanogr. 33(2):465–477.

Hedrick, P.W.

1999. Perspective: highly variable loci and their interpretation in evolution and conservation. Evolution 53, 313–318.

Jakobsson, M., and N. A. Rosenberg.

2007. CLUMPP: a cluster matching and permutation program for Dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23(14):1801–1806.

Kalinowski, S.T.

2002. How many alleles per locus should be used to estimate genetic distances? Heredity 88, 62–65.

Kalinowski, S.T.

2005. Do polymorphic loci require large sample sizes to estimate genetic distances? Heredity 94:33–36.

Kingsley, M. C. S.

2004. The hogfish in Florida: Assessment Review and Advisory Report. Pages 15 *in* Southeast Data and Assessment Review. Prepared for the South Atlantic Fishery

Management Council, the Gulf of Mexico Fishery Management Council, and the National Marine Fisheries Service.

Latter, B.D.H.

1972. Selection in finite populations with multiple alleles. III. genetic divergence with centripetal selection and mutation. Genetics 70:475–490.

Lindholm, J., A. Knight, L. Kaufman, and S. Miller.

2006. A pilot study of hogfish (*Lachnolaimus maximus* Walbaum 1792) movement in the Conch Reef Research Only Area (Northern Florida Keys). Marine Sanctuaries Conservation Series NMSP–06–06.

Lischer H.E.L, and L. Excoffier.

2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28:298–299.

Liu, Y. G., and R. H. Weisberg.

2005. Patterns of ocean current variability on the West Florida Shelf using the self-organizing map. J. Geohys. Res., C 110(C6).

Liu, Y. G., and R. H. Weisberg.

2012. Seasonal variability on the West Florida Shelf. Prog. Oceanogr. 104:80–98.

Mariani, S., and D. Bekkevold.

2014. The nuclear genome: Netural and adaptive markers in fisheries science. Pages 297–320 *in* S. X. Cadrin, L. A. Kerr, and S. Mariani, editors. Stock Identification Methods (Second Edition). Academic Press, San Diego. Mol. Ecol. Res. Primer Development Consortium.

Malecot, G.

1955. Decrease of relationship with distance. Cold Spring Harbor Syrup. Quant Biol 20:52–53.

McBride, R. S.
in press. Managing a marine stock portfolio: stock identification, structure, and management of 25 fishery species along the Atlantic coast of the United States. North Am. J. Fish. Manage.

McBride, R. S.
2014. The continuing role of life history parameters to identify stock structure. Pages 77–107 *in* S. X. Cadrin, L. A. Kerr, and S. Mariani, editors. Stock Identification Methods (Second Edition). Academic Press, San Diego.

McBride, R. S., and M. R. Johnson.
2007. Sexual development and reproductive seasonality of hogfish (Labridae: *Lachnolaimus maximus*), an hermaphroditic reef fish. J. Fish Biol. 71(5):1270–1292.

McBride, R. S., and M. D. Murphy.
2003. Current and potential yield per recruit of hogfish, *Lachnolaimus maximus*, in Florida. Proceedings of the Gulf and Caribbean Fisheries Institute 54:513–525.

McBride, R. S., and A. K. Richardson.
2007. Evidence of size-selective fishing mortality from an age and growth study of hogfish (Labridae: *Lachnolaimus maximus*), a hermaphroditic reef fish. Bull. Mar. Sci. 80(2):401–417.

McBride, R. S., P. E. Thurman, and L. H. Bullock.
2008. Regional variations of hogfish (*Lachnolaimus maximus*) life history: consequences for spawning biomass and egg production models. J. Northwest Atl. Fish. Sci. 41:1–12.

Muñoz, R. C., M. L. Burton, K. J. Brennan, and R. O. Parker.
2010. Reproduction, habitat utilization, and movements of hogfish (*Lachnolaimus maximus*) in the Florida Keys, U.S.A.: comparisons from fished versus unfished

habitats. Bull. Mar. Sci. 86(1):93–116.

NMFS (National Marine Fisheries Service).
2013. Commercial Fisheries Statistics.
(http://www.st.nmfs.noaa.gov/st1/commercial/index.html).

Paris, C. B., J. Atema, J.-O. Irisson, M. Kingsford, G. Gerlach, and C. M. Guigand.
2013. Reef Odor: A Wake Up Call for Navigation in Reef Fish Larvae. Plos One
8(8):e72808.

Park, S. D. E.
2001. Trypanotolerance in West African Cattle and the Population Genetic Effects of
Selection. Ph.D. thesis, 241 p. University of Dublin, Republic of Ireland.

Parker, R.O.
2000. Courtship in hogfish, *Lachnolaimus maximus*, and other behavior of reef
fishes off Beaufort, North Carolina. J. Elisha Mitchell Sci. So. 116 (3):260–261.

Peakall, R., and P.E. Smouse.
2012. GenAlEx (version 6.5): genetic analysis in Excel. Population
genetic software for teaching and research-an update. Bioinformatics 28:2537–2539.

Peakall, R., and P.E. Smouse.
2006. GENALEX (version6): Genetic analysis in Excel. Population
genetic software for teaching and research. Mol. Ecol. Notes 6:288–295.

Pritchard, J. K., M. Stephens, And P. Donnelly.
2000. Inference of population structure using multilocus genotype data. Genetics
155:945–959.

Pritchard, J. K., X. Wen, And D. Falush.
2009. Documentation for structure software (version 2.2). University of Chicago.
Available: pritch.bsd.uchicago.edu/software.

586 Purcell, J. F. H., R. K. Cowen, C. R. Hughes, and D. A. Williams.
587 2006. Weak genetic structure indicates strong dispersal limits: a tale of two coral reef
588 fish. Proceedings of the Royal Society B: Biological Sciences 273(1593):1483–1490.
589 Raymond, M., and F. Rousset.
590 1995. An exact test for population differentiation. Evolution 49:1280–1283.
591 Rice, W.R.
592 1989. Analyzing tables of statistical tests. Evolution 43:223–225.
593 Richards, W. J., M. F. McGowan, T. Leming, J. T. Lamkin, and S. Kelley.
594 1993. Larval fish assemblages at the loop current boundary in the Gulf of Mexico.
595 Bull. Mar. Sci. 53(2):475–537.
596 Rousset, F.
597 2008. Genepop'007: a complete reimplementation of the GenePop software for
598 Windows and Linux. Mol. Ecol. Res. 8:103–106.
599 Rosenberg, N.A.
600 2004. DISTRUCT: a program for the graphical display of population
601 structure. Mol. Ecol. Notes 4:137–138.
602 Seyoum, S., M.D. Tringali, B.L. Barthel, C. Puchulutegui, M.C. Davis, A.B. Collins, and
603 R.S. McBride.
604 2012. Twenty nine polymorphic microsatellite loci for the genetic management of
605 captive bred hogfish (*Lachnolaimus maximus*) and assessments of globally declining
606 populations. (Permanent genetic resources added to Mol. Ecol. Res. Database,
607 February 1, 2012-March 31, 2012). Mol. Ecol. Res. 12:779–781.
608 http://biomath.Trinity.edu/manuscripts/12-4/mer-12–0057.pdf
609 Seyoum, S., M.D. Tringali, T.M. Bert, D. McElroy, and R. Stokes.

2000. An analysis of genetic population structure in red drum, *Sciaenops ocellatus*, based on mtDNA control region sequences. Fishery Bulletin U. S. 98(1):127–138.

Seyoum, S., M.D. Tringali, B.L. Barthel, V. Villanova, C. Puchulutegui, M.C. Davis, and A. C.C. Alvarez.

2014. Stock boundaries for spotted seatrout (*Cynoscion nebulosus*) in Florida based on population genetic structure. Fish and Wildlife Research Institute Technical Report XX–Y. ooo + 000 p.

Soeparno, Y. Nakamura, T. Shibuno, and K. Yamaoka.

2012. Relationship between pelagic larval duration and abundance of tropical fishes on temperate coasts of Japan. J. Fish Biol. 80(2):346–357.

Takezaki, N., M. Nei, and K. Tamura.

2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows-interface. Molecular Biology and Evolution 27:747–752.

Tringali, M. D., and T. M. Bert.

1996. The genetic stock structure of common snook (*Centropomus undecimalis*). Can. J. Fish. Aquat. Sci. 53:974–984.

Tringali, M. D., and M. Higham.

2007. Isolation-by-Distance Gene Flow among Vermilion Snapper (*Rhomboplites aurorubens* Cuvier, 1892) from the Gulf of Mexico and Southeastern United States. Gulf of Mexico Science 2007(1):2–14.

Victor, B. C.

1986. Duration of the planktonic larval stage of one hundred species of Pacific and Atlantic wrasses (family Labridae). Mar. Biol. 90:317–326.

Waples, R. S.

1987. A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution 41(2):385–400.

Weir, B. S., and C.C. Cockerham.

1984. Estimating F-statistic for the analysis of population structure. Evolution 38:1358–1370.

Wilson, G.A., and B. Rannala.

2003. Bayesian inference of recent migration rates using multilocus genotypes. Genetics, 163:1177–1191.

Wright, S.

1943. Isolation by distance. Genetics 28(2)114–138.

Zatcoff, M. S., A. O. Ball, and G. R. Sedberry.

2004. Population genetic analysis of red grouper, *Epinephelus morio*, and scamp, *Mycteroperca phenax*, from the southeastern US Atlantic and Gulf of Mexico. Mar. Biol. 144(4):769–777.

Figure 1. Hogfish (*Lachnolaimus maximus*) specimens were grouped by collection location into samples, where 1 = Big Bend, 2 = Nature Coast and Middle Grounds, 3 = Tampa Bay, 4 = Sarasota, 5 = Naples, 6 = Everglades, 7 = Florida Keys, 8 = East Florida, 9 = Carolinas.

Figure 2. Phenograms based on (A) uncorrected and (B) corrected genetic distances (F_{ST}) between each pair of the 9 hogfish samples as identified through the Unweighted Pair Group Method with Arithmetic Mean (UPGMA: 1000 bootstrap replications) hierarchical clustering method using the program POPTREE2.

Figure 3. (A) Mean likelihood L (posterior probability) and (B) ΔK as determined from 10 replicates of each value of K (from 1 to 9). Statistics are from STRUCTURE analysis of the 9 hogfish collection samples.

Figure 4. Genetic population structure among hogfish samples according to posterior probability assignment produced by the STRUCTURE analysis of 24 polymorphic microsatellite loci (4.5×10^6 burn-in, and 9.0×10^6 replications): The CLUMMP output from 10 replicates for 719 individuals and the 9 samples; (A) for the highest modal value (number of genetic clusters) at K=2, (B), for the next modal value of K=3 shown in graphic display (hogfish population structure) using the program DISTRUCT, and (C) proportional values of each cluster for each of the 9 samples. Figures are demarcated by vertical lines.

Figure 5. Mantel tests examining correlations between pairwise genetic distance (F_{ST}) and geographic distance (km) within Cluster 1 from Big Bend (sample 1) to the Everglades (sample 6).

Figure 6. Geographic boundaries of the three hogfish distinct clusters based on the genetic population structure, as detected using 24 microsatellite loci. Dotted lines indicate flexible demarcations of clusters potentially expanding in the direction indicated by arrows. The region of gene flow restriction (genetic break) between clusters 1 and 2 is identified by a circular delineation.

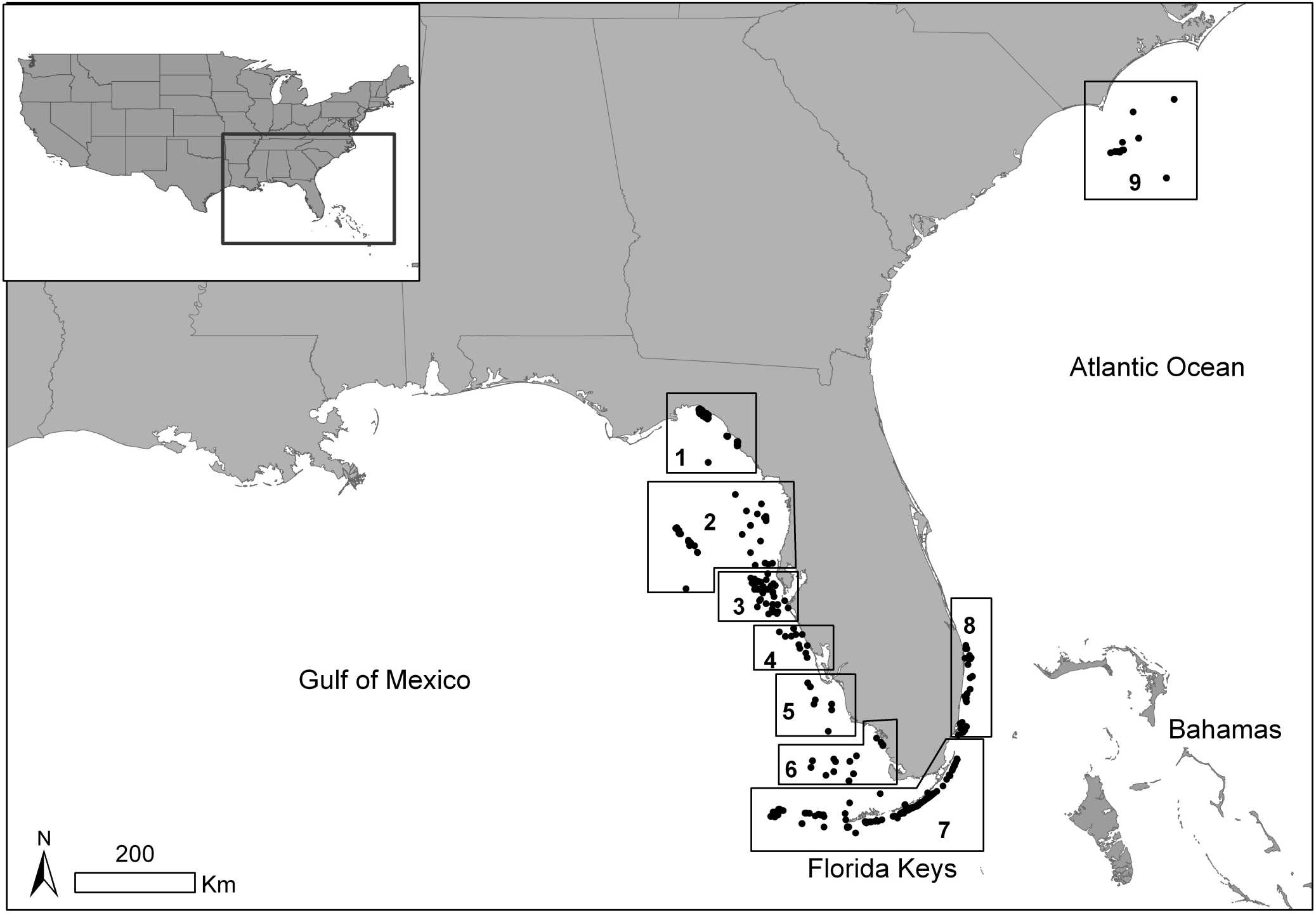
Table 1. Standard genetic measurements averaged over all loci for each of the nine samples of hogfish. Collection areas for samples appear in Figure 1. In the last column, numbers in parenthesis represent sample averages.

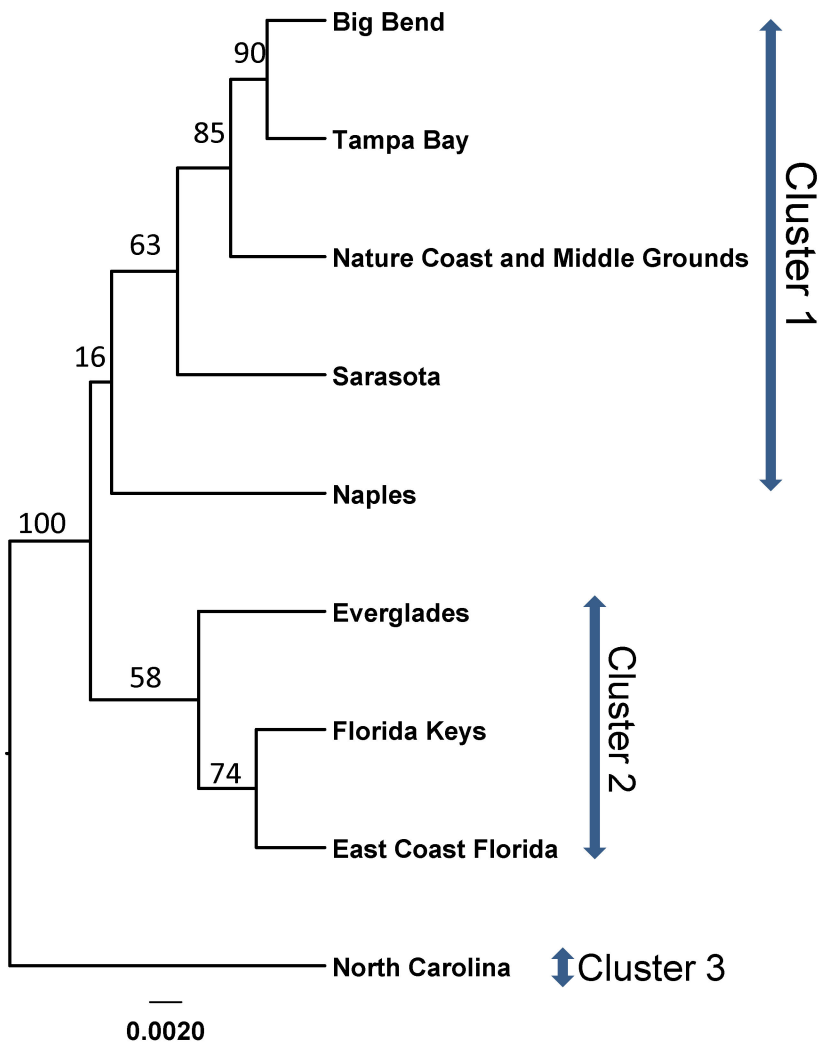
	Samples									
	1	2	3	4	5	6	7	8	9	Total
No. of specimens	119	71	88	24	22	70	191	32	102	719
Genetic diversity	0.63	0.65	0.65	0.63	0.64	0.67	0.67	0.68	0.61	0.65 (0.65)
Total no. of alleles sampled	246	223	237	165	161	238	296	203	174	350 (216)
Average no. of alleles per locus	10.3	9.3	9.9	6.9	6.7	9.9	12.3	8.5	7.3	14.6 (9.0)
Allelic richness	9.5	9.0	9.7	9.6	6.3	9.7	12.2	8.2	7.1	14.4 (9.0)
Observed heterozygosity	0.57	0.59	0.60	0.61	0.59	0.63	0.63	0.63	0.58	0.60 (0.60)
Unbiased observed heterozygosity	0.63	0.65	0.65	0.63	0.64	0.66	0.67	0.68	0.61	0.66 (0.65)
Expected heterozygosity	0.63	0.64	0.64	0.62	0.63	0.66	0.66	0.67	0.60	0.66 (0.64)

Table 2. Analysis of molecular variance (AMOVA) based on 24 microsatellite loci genotypes of 719 hogfish specimens collected from Big Bend, Florida, to the Carolinas. Numbers joined by hyphens are samples combined into clusters based on the corrected and uncorrected F_{ST} distances. Number codes as in Figure 1. Only the values of the F_{CT} among the three-cluster and the F_{ST} between the two-cluster analyses are given below.

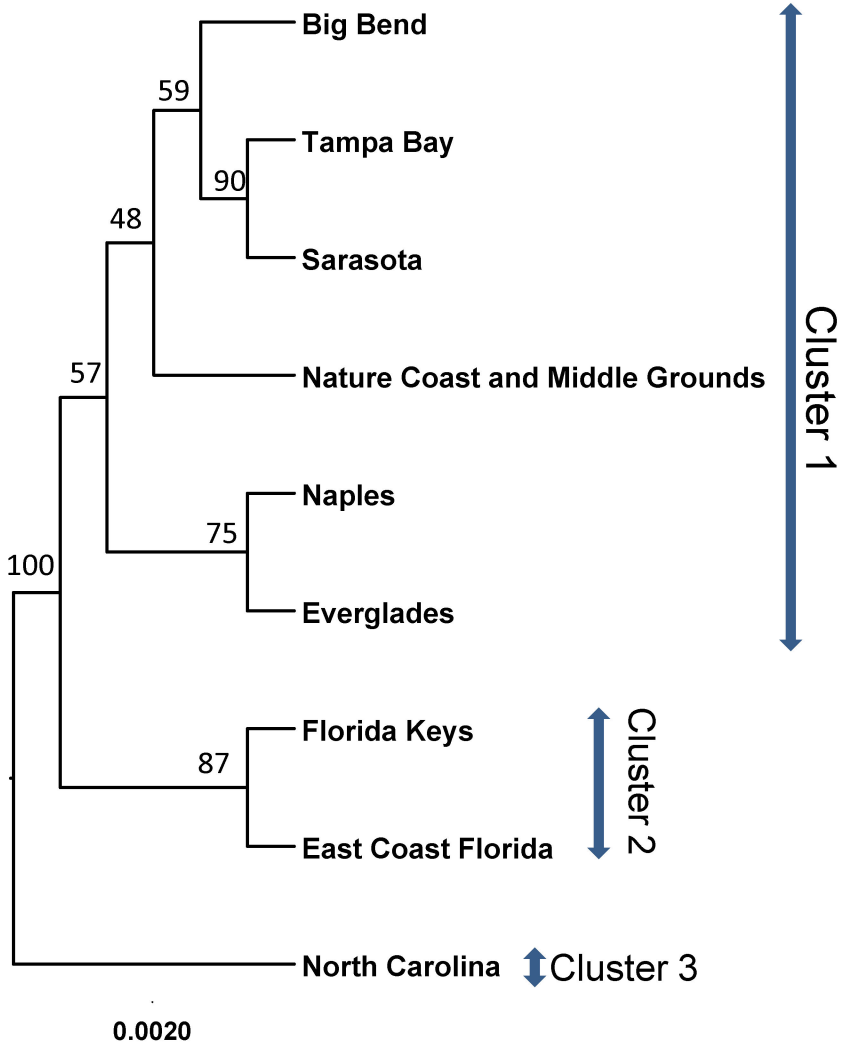
Among three-cluster component			Observed partition			
1st Cluster	2nd Cluster	3rd Cluster	Variance	% Total	F_{CT}	Phenetic cluster from
1-2-3-4-5-6	7-8	9	0.14768	2.07	0.0208	Corrected F_{ST}
1-2-3-4-5	6-7-8	9	0.14331	2.02	0.0202	Uncorrected F_{ST}
Among two-cluster component			**Observed partition**			
1st Cluster	2nd Cluster	3rd Cluster	Variance	% Total	F_{ST}	Phenetic cluster from
1-2-3-4-5-6	7-8	-----	0.11203	1.56	0.0156	Corrected F_{ST}
1-2-3-4-5-6	-----	9	0.16073	2.30	0.0230	Corrected F_{ST}
-----	7-8	9	0.28595	3.93	0.0393	Corrected F_{ST}
1-2-3-4-5	6-7-8	-----	0.10629	1.48	0.0149	Uncorrected F_{ST}
1-2-3-4-5	-----	9	0.17683	2.56	0.0256	Uncorrected F_{ST}
-----	6-7-8	9	0.22038	3.01	0.0301	Uncorrected F_{ST}

Probability of finding a more extreme variance component by chance alone (1000 permutations) <0.001 in all the F_{CT} and F_{ST} statistic.

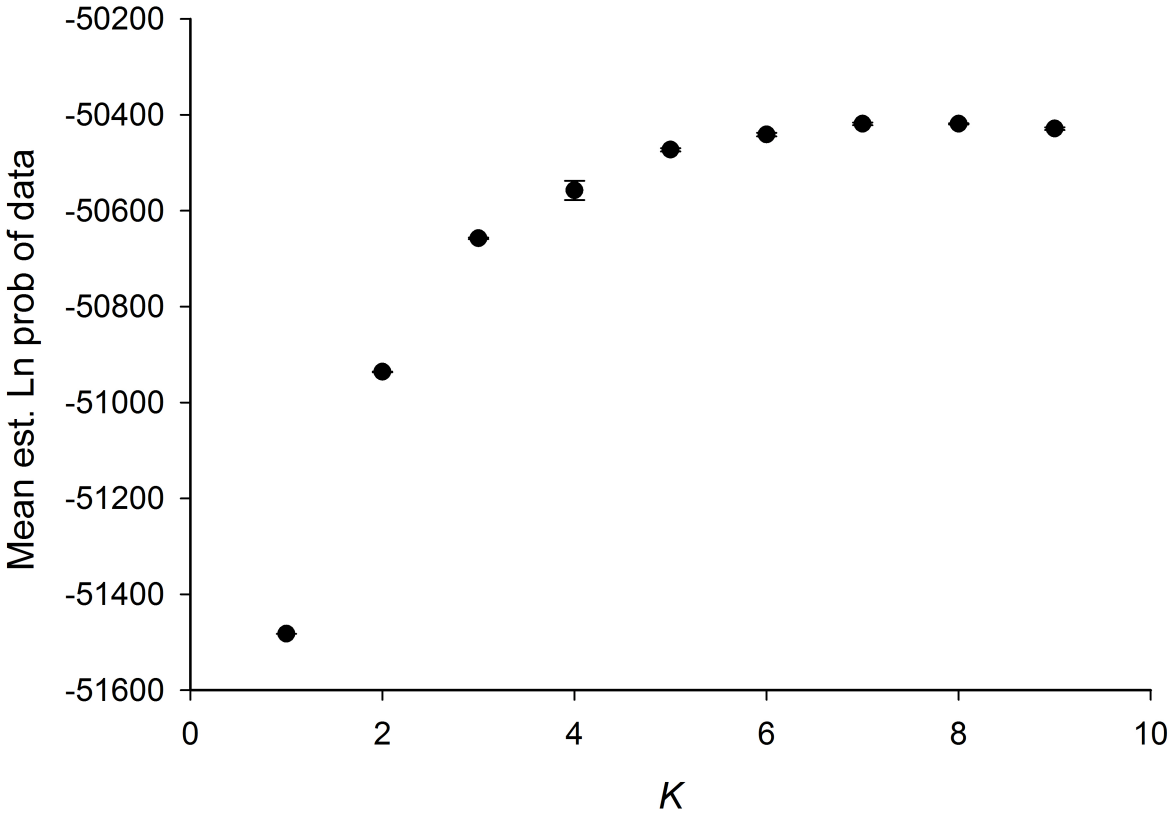

Atlantic Ocean
Gulf of Mexico
Bahamas
Florida Keys
1
2
3
4
5
6
7
8
9
N
200
Km

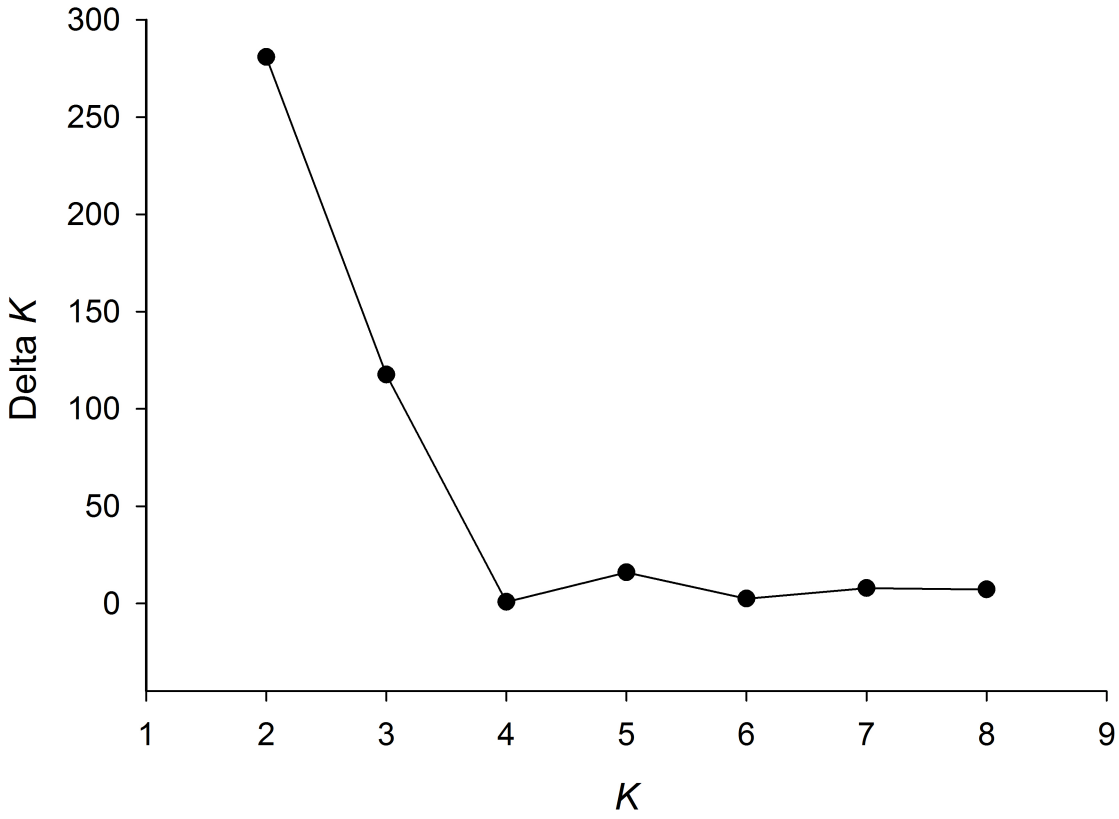


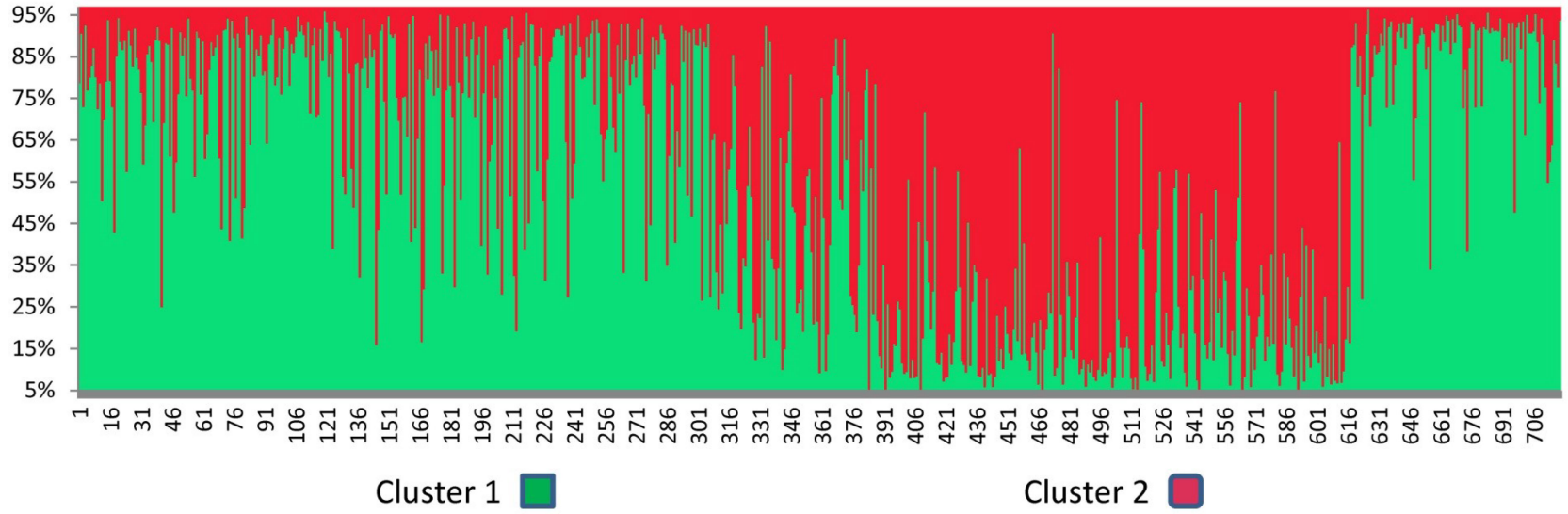
(A) Uncorrected F_{ST}

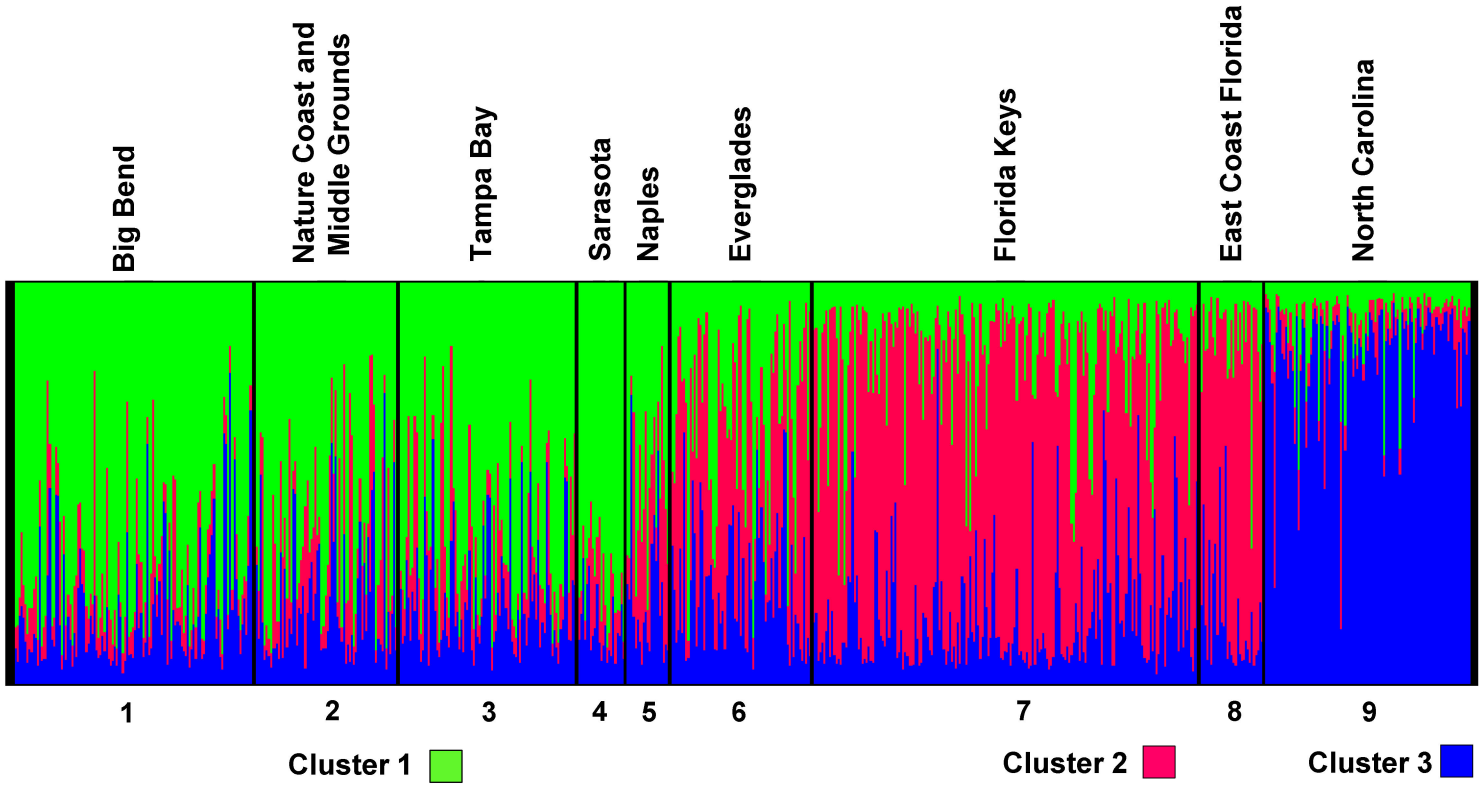


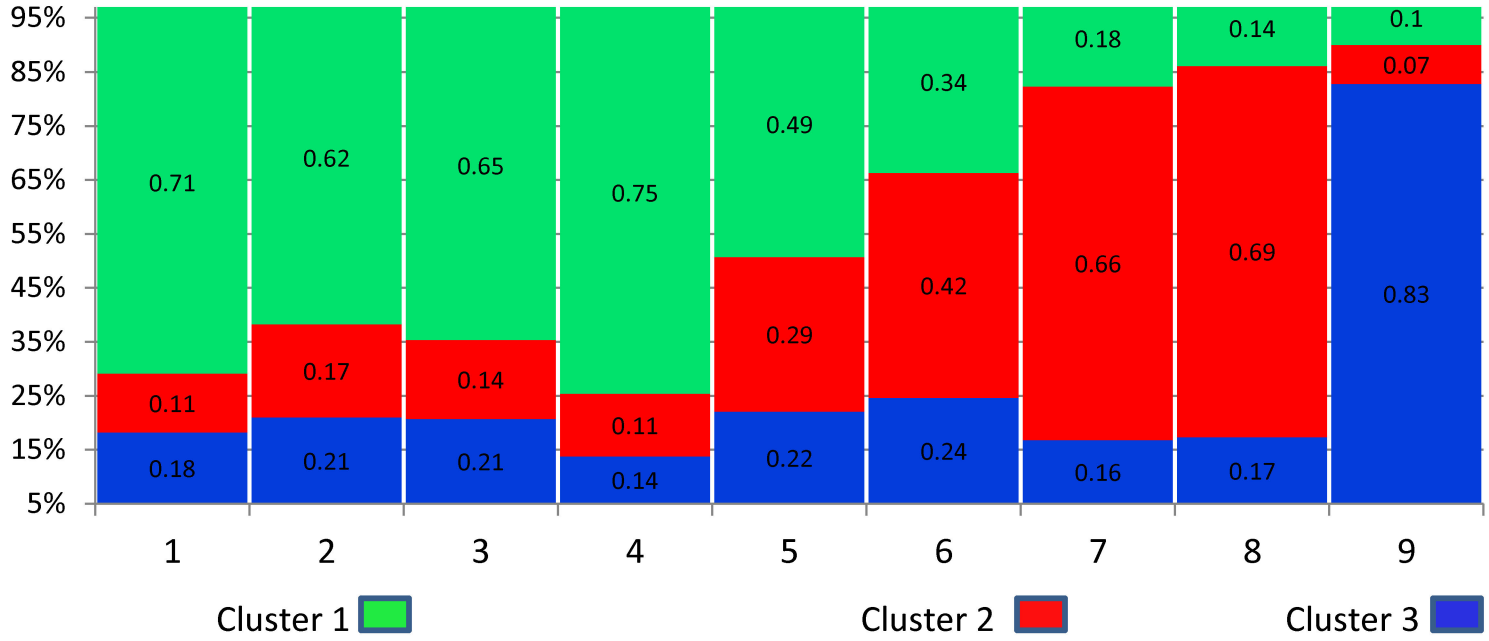
(B) Corrected F_{ST}

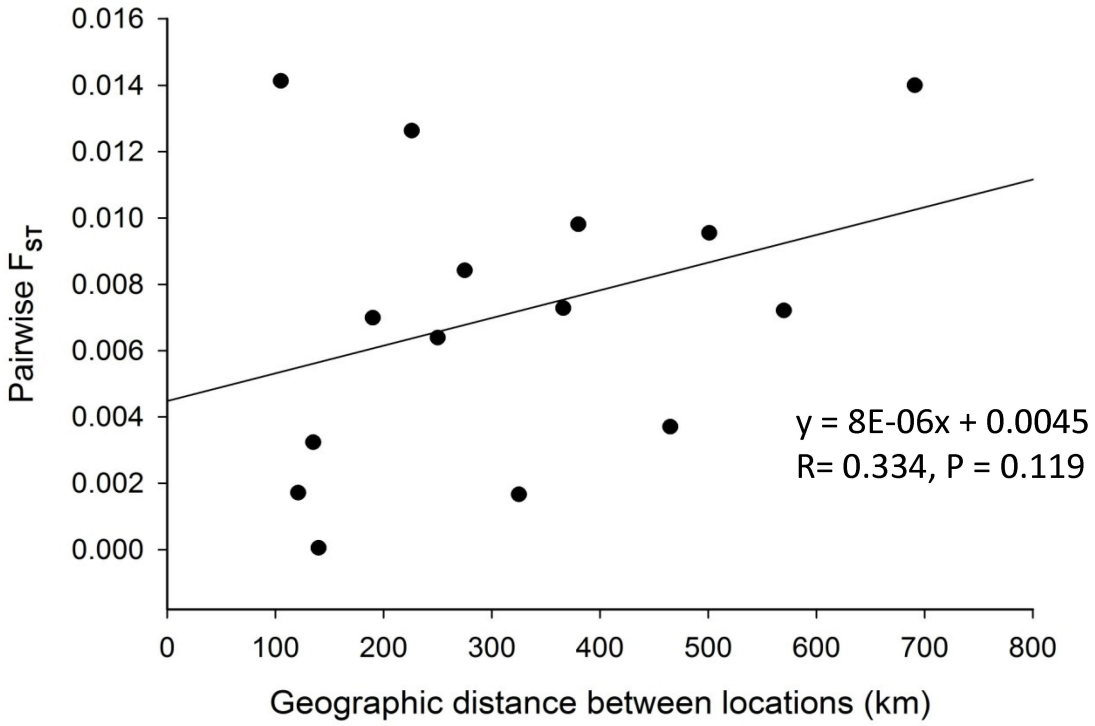

-50200
-50400
-50600
-50800
-51000
-51200
-51400
-51600
Mean est. Ln prob of data
0
2
4
6
8
10
K

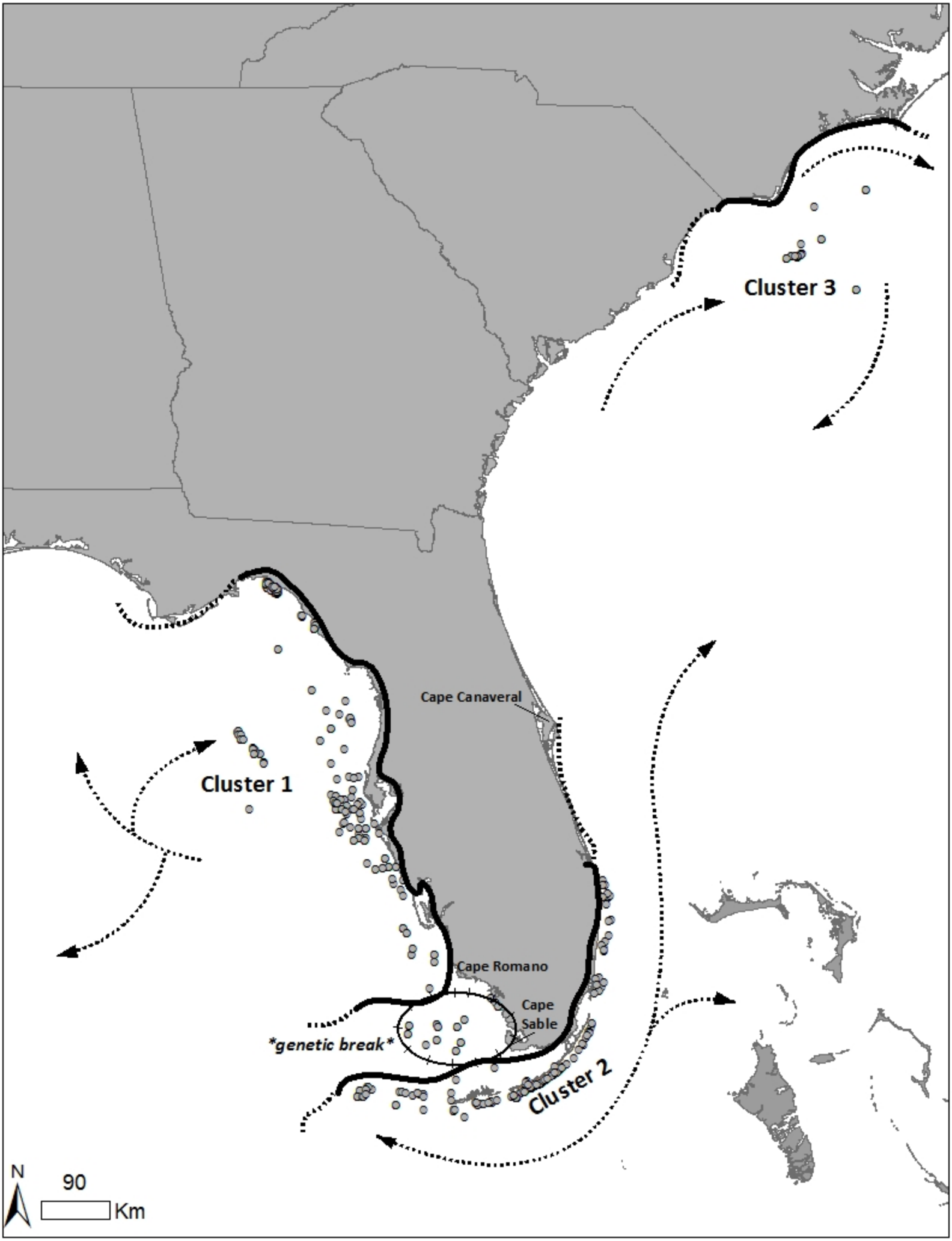

300
250
200
150
100
50
0
Delta K
1
2
3
4
5
6
7
8
9
K


95%
85%
75%
65%
55%
45%
35%
25%
15%
5%
1
16
31
46
61
76
91
106
121
136
151
166
181
196
211
226
241
256
271
286
301
316
331
346
361
376
391
406
421
436
451
466
481
496
511
526
541
556
571
586
601
616
631
646
661
676
691
706
Cluster 1
Cluster 2


Big Bend
Nature Coast and Middle Grounds
Tampa Bay
Sarasota
Naples
Everglades
Florida Keys
East Coast Florida
North Carolina
1
2
3
4
5
6
7
8
9
Cluster 1
Cluster 2
Cluster 3


95%
85%
75%
65%
55%
45%
35%
25%
15%
5%
0.71
0.11
0.18
0.62
0.17
0.21
0.65
0.14
0.21
0.75
0.11
0.14
0.49
0.29
0.22
0.34
0.42
0.24
0.18
0.66
0.16
0.14
0.69
0.17
0.1
0.07
0.83
1
2
3
4
5
6
7
8
9
Cluster 1
Cluster 2
Cluster 3


0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
Pairwise F_{ST}
0
100
200
300
400
500
600
700
800
Geographic distance between locations (km)
y = 8E-06x + 0.0045
R= 0.334, P = 0.119


Cluster 3
Cape Canaveral
Cluster 1
Cape Romano
Cape
Sable
genetic break
Cluster 2
N
90
Km