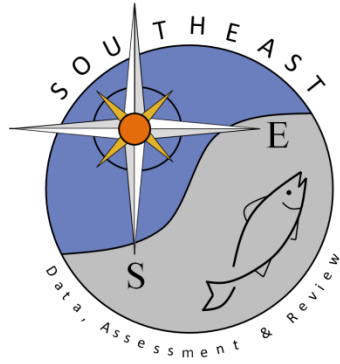


Assessment of Genetic Stock Structure of Gray Triggerfish (*Balistes capricus*) in U.S. Waters of
the Gulf of Mexico and South Atlantic Regions

Eric Saillant and Luca Antoni

SEDAR41-RD37

23 July 2014



Final Report

**ASSESSMENT OF GENETIC STOCK STRUCTURE OF GRAY
TRIGGERFISH (*BALISTES CAPRISCUS*) IN U.S. WATERS OF
THE GULF OF MEXICO AND SOUTH ATLANTIC REGIONS**

MARFIN award # NA09NMF4330150

Eric Saillant, Luca Antoni

Department of Coastal Sciences, The University of Southern Mississippi, Gulf Coast Research
Laboratory, Ocean Springs, MS, 30564

March 31 2014

Executive summary

The gray triggerfish (*Balistes capriscus*) is a reef fish exploited by recreational and commercial fisheries in the Southeast United States. Recent stock assessments indicate that the species is overfished and experiencing overfishing. A fisheries management plan was initiated in 2008 and aims to end overfishing and rebuild the stock. Currently the gray triggerfish resource is managed under the jurisdictions of the Gulf of Mexico and South Atlantic regions with no subdivision in geographic stocks. The objectives of this project were to develop molecular tools for genetic assessment and monitoring of gray triggerfish, and provide a first assessment of genetic stock structure in U.S. waters of the northern Gulf of Mexico and South Atlantic regions where the species is exploited.

Twenty one homologous microsatellites were developed from enriched genomic libraries. Seventeen of the microsatellites were combined in multiplex panels and used in studies of genetic variation along with a sequencing assay of a 617 bp fragment of the mitochondrial coding gene ND4. The markers were tested in two other economically important balistids, the queen triggerfish (*Balistes vetula*) and the ocean triggerfish (*Canthidermis sufflamen*). Fourteen microsatellites were found usable for genetic studies of the queen triggerfish while only four were suitable for studies of ocean triggerfish.

A total of 665 gray triggerfish were collected in U.S. waters including specimens from 6 localities in the Gulf of Mexico and 2 localities in the South Atlantic region (61-80 samples per locality); 68 additional specimens from the northern Gulf of Mexico were included in order to better reflect the continuous distribution of gray triggerfish in that region. Analysis of genetic variation at the 17 microsatellites and mitochondrial DNA revealed homogeneous distributions of allele and haplotype frequencies throughout the sampling surface. Spatial analyses and Bayesian clustering failed to detect occurrence of genetic discontinuities or cryptic structuring within the dataset. Occurrence of isolation-by-distance was investigated using a regression approach. Estimates of the slope of the isolation-by-distance model were generated based on individual data obtained in the northern Gulf of Mexico where sampling reflected best the continuous distribution of gray triggerfish on shelf habitats. The estimated slope (point estimate 3.14×10^{-8} upper bound 4.61×10^{-7}) was consistent with large neighborhood sizes. Dispersal distributions were simulated based on three different families of distribution functions with parameters adjusted to yield isolation-by-distance slopes similar to that obtained with the empirical dataset. The simulated functions involved mean dispersal distances greater than 237 km and substantial fractions of dispersal events at large distances (greater than 500 km). Improved estimates of population density and studies of temporal stability are warranted in order to obtain more accurate estimates of the dispersal distribution. However, the inference of large dispersal distances is consistent with the high dispersal potential of larvae and juveniles through transport by the *Sargassum* and suggests that regional stocks are partially dependent demographically.

Connectivity of U.S. stocks with populations in the East Atlantic was investigated through analysis of genetic variation in 64 specimens from France and 72 specimens from Benin. The distribution of mitochondrial DNA haplotypes and microsatellite alleles were homogeneous between the U.S. and French samples suggesting high connectivity that is hypothesized to reflect transport of juveniles by the Gulf Stream. In contrast, the sample from Benin showed pronounced divergence in allele frequencies with both U.S. and French populations (F_{ST} estimates were 0.17 at microsatellites and 0.30 at mitochondrial DNA) suggesting isolation

between West Africa and both the West Atlantic and European populations. Restricted gene flow between these groups is consistent with the limited apparent surface circulation connecting these regions.

The suggestion of long distance connectivity through dispersal of larvae suggests that breeding stocks in the Caribbean and South America might contribute to recruitment in U.S. waters. Studies of the connectivity between U.S. populations and those regions are warranted.

1. Purpose

The main purpose of this project is to document genetic stock structure and connectivity in an exploited reef fish, the gray triggerfish *Balistes capriscus*, to assist in allocation, assessment, and management of U.S. stocks. Knowledge of stock structure is essential to develop effective management plans for fisheries. Failure to recognize existing sub units of an exploited stock may lead to overexploitation and depletion of local populations (Carvalho and Hauser 1994, Begg et al. 1999), lack of success of management measures (Stephenson 1999), and inability to anticipate recruitment in regional fisheries during the course of rebuilding programs (Ruzzante et al. 1999). An additional long term consequence of local depletion is that unique genetic characteristics of local stocks can be lost leading to loss of biodiversity at the metapopulation level (Waples 1995). Unique genetic characters found in local populations are involved in local adaptation and health and fitness of local stocks but also contribute to the metapopulation long term evolutionary potential and its capability to sustain environmental changes (Hilborn et al. 2003). Genetic stock assessment and conservation of genetically distinct local populations should therefore be a primary focus in order to maximize evolutionary potential and sustainability of stocks. Development of a reliable model for genetic assessment of fisheries stocks requires deployment of molecular markers and population genetics models which allow inferring population/stock genetic structure, along with some aspects of stock demography including genetic migrations or gene flow among geographic populations and population expansion or decline (Beaumont 2001, Allendorf and Luikart 2007).

The gray triggerfish is a reef fish commonly found in temperate and tropical shelf waters of the Gulf of Mexico and Western Atlantic basins. In the United States, the species is exploited by commercial and recreational fisheries, primarily in the northern Gulf of Mexico and South Atlantic regions. Most of the landings are made by recreational fisheries catching gray triggerfish using handlines but the species is also landed by commercial fishers using various gears including handlines, longlines, traps and trawls (Goodyear and Thompson 1993). Gray triggerfish was historically not considered a desirable catch but the recent decline and increased regulations of other reef species such as the red snapper have resulted in increased fishing efforts towards this species. Accordingly, landings increased during the late 80's from 1.46 Million lbs to reach 2.88 Million lbs in 1990 (Valle et al. 2001) but then regularly declined beginning in the early 90's. A stock assessment conducted in 2001 (Valle et al. 2001) indicated that the stock was likely overfished and that mortality rates were non sustainable. The assessment conducted in 2006 (SEDAR 9) concluded that the species was not overfished but status was subsequently re-evaluated accounting for higher minimum benchmarks for spawning potential ratios of stocks. Recent data for the Gulf of Mexico and South Atlantic regions showed further decrease of gray triggerfish catches from 87,168 lbs in 2004 to 8,769 lbs in 2012 (commercial catches) and from

1,364,413 lbs to 519,293 lbs over the same period for the recreational fishery (Personal communication from the National Marine Fisheries Service, Fisheries Statistics Division). Because the development of gray triggerfish fisheries is relatively recent, most of the efforts to assess and manage this species were conducted during the past decade. Gray triggerfish has been managed under the 1981 Reef Fishery Management Plan (RFMP) implemented in 1984 which provided a list of species and defined an inshore area in which certain gears were prohibited. The RFMP amendment 12 established an aggregate bag limit of 20 fish for all listed species (including gray triggerfish) but with no per-species bag limit. A minimum size of 12 inches was established by Amendment 16B in 1999. A new regulation was published July 3 2008 (<http://www.gpoaccess.gov/fr/index.html>) and established a new minimum size of 14 inches, annual catch limits and accountability measures that use average annual landings for the recreational sector and in-season adjustments for the commercial sector. The new regulation also established a commercial quota of 80,000 pounds for 2008, 93,000 pounds for 2009, and 106,000 pounds for 2010 (NOAA Fisheries southeast fishery bulletin FB08-40, July 2008) for Gulf of Mexico gray triggerfish. Following the assessment update conducted in 2011, the species has been under annual catch limits for commercial and recreational fisheries in both the South Atlantic and Gulf of Mexico regions.

The Southeast Data Assessment and Review (SEDAR) 9 on gray triggerfish put a strong emphasis on defining management criteria and stock benchmarks. Specifically, various indices and analytical models were examined for use in assessment, and possible rebuilding trajectories were evaluated. The validity of considering multiple geographic sub-units of the stock was left un-clarified.

Gray triggerfish are currently managed under the jurisdiction of the Regional Fisheries Management Councils (primarily Gulf of Mexico and South Atlantic regions). Management thus makes the implicit assumption that the species constitutes a single stock within each region (Gulland 1965, Hilborn and Walters 2001) and that these two regional stocks are demographically independent. Occurrence of multiple stocks (e.g. northeastern versus northwestern Gulf of Mexico) is however a possibility that remained to be tested rigorously at the beginning of this project. Similarly, considering the high dispersal potential of gray triggerfish at early life stages (see below), partial demographic dependence of the South Atlantic and Gulf of Mexico populations was a possibility that deserved to be examined. Previous assessments of gray triggerfish in the Gulf of Mexico highlighted occurrence of erratic variation in life history and phenotypic traits over small geographic scales. Because these variations potentially reflect, in part, effects of differing fishing pressures among sub-regions and/or phenotypic plasticity, the definition of multiple stocks on the basis of phenotypic and/or life history traits appears problematic.

The life history of gray triggerfish is characterized by a high potential for dispersal at early life stages. Adult gray triggerfish are usually found in proximity of bottom structures at depths between 10 and 106 m. They spawn in demersal nests guarded by the male parent. After hatch, larvae become pelagic (Richards and Linderman 1987) and juveniles are found associated with seaweed and flotsam (mostly *Sargassum* sp.) where they can stay for a relatively long time (up to 3.15 in SL according to Wells and Rooker 2004, and up to 5 in in the study by Franks et al. 2007). The early life history of gray triggerfish thus indicates high dispersal potential during their early larval phase and/or via subsequent passive transport during association with *Sargassum* beds. Juveniles are found settling on benthic habitat when they reach 16-64 in

(Ingram 2001) at an age of 4 to 7 months (Simmons and Szedlmayer 2011). Following settlement, tag and recapture studies indicate extremely limited movement potentially owing to territorial behaviors (Ingram 2001). Longevity is intermediate with most individuals found between 2 and 6 years (Wilson et al. 1995). Gray triggerfish behavioral and life history characteristics potentially make them vulnerable to overfishing because of (i) a limited potential for re-colonization of habitats by adults following overfishing and (ii) potential significant modification of the age and size structure of local populations by fishing activities considering that captures impact preferentially large males that display aggressive behavior. The latter has a potential genetic impact in that systematical removal of larger fish could lead to artificial selection of slow growing fish (Conover et al. 2005) which could be deleterious if slow growing fish have lower fitness, e.g. lower reproductive potential or other negative correlated responses to this selective pressure (Reznick and Ghalambor 2005).

In consequence, inference on stock structure based on life history or phenotypic traits (e.g. growth rate) are particularly unsuited in gray triggerfish as growth characteristics of local populations may reflect primarily the degree of fishing pressure in that population. Indeed, observations of large variations in growth over small spatial scales are frequent and are thought to likely reflect of different fishing pressures among regions (Ingram 2001).

In summary, potential for dispersal of gray triggerfish is high during the early life phases especially if the *Sargassum* or other floatsam habitats used by juveniles are transported over long distances by oceanic currents. However, cyclonic or anti-cyclonic currents may retain most young triggerfish in the area where they were spawned leading to mostly self-recruited local populations (Ingram 2001, SEDAR #9, Gulf of Mexico Gray Triggerfish, Assessment report 1, 2006). Overall, considering that the time frame during which dispersal occurs is limited to 4-7 months of pelagic transport, an isolation-by-distance model of population structure is expected where genetic distance increases as a function of geographic distance reflecting limited dispersal. A rigorous examination of genetic stock structure appears therefore especially relevant considering the limits of comparisons of life history and phenotypic traits highlighted above.

This project contributes to filling a major gap in knowledge of stock structure and connectivity of gray triggerfish by examining patterns of genetic variation among geographic. At the beginning of this project there was no molecular marker available for gray triggerfish. The first phase of this work therefore focused on developing a robust panel of markers for genetic monitoring. The markers were then deployed to provide a first assessment of genetic stock structure in U.S waters. Specifically, the questions addressed by the project include (i) are there multiple genetic stocks of gray triggerfish in the southeastern United States where the species is exploited and what are the relationships (genetic migrations or gene flow) among those units if they exist (ii) are the data consistent with an isolation-by-distance model of population structure (iii) are there differences in (genetic) effective size and genetic diversity among sub units of the stock and is there a signature of a reduction in effective population size (genetic bottleneck) in the region in relation to decline in abundance and potential overfishing.

An additional question that arose during the course of the project and considering both the high dispersal potential of larvae and the homogeneity in allele frequencies among U.S. populations is that of the connectivity of U.S. stocks with foreign stocks. Although beyond the scope of the initial proposal, we initiated research to study connectivity of U.S. triggerfish with other populations within the distribution range.

Specific project objectives are therefore as below

Objective 1. Develop a robust panel of molecular markers for genetic monitoring of gray triggerfish. Optimize assays and characterize markers for use in genetic studies of gray triggerfish and two other economically important triggerfish species (the queen triggerfish and the ocean triggerfish)

Objective 2. Survey genetic variation at the developed molecular markers in geographic populations within U.S. waters and analyze data to characterize stock structure and connectivity.

Objective 3. Initiate study of the connectivity of U.S. stocks with foreign stocks

2. Approach

2.1 Approach objective 1.

The objective of this project section was to develop a panel of 13-15 microsatellites suitable for population genetics studies of gray triggerfish and compatible for high throughput assay in multiplex. Our experience with other marine species including reef fishes is that such a panel of microsatellites achieves high power of inference for analyses of stock structure and demographic parameters. A sequencing assay for a coding fragment of mitochondrial DNA (mtDNA) was also developed. While mtDNA only represents one locus, its haploid nature and maternal inheritance leads to increased effects of genetic drift, potentially increasing resolution power relative to nuclear markers. The main interest in mtDNA markers is for use in comparison with nuclear markers as this allows examination, at least qualitatively, of differences between sexes in dispersal and genetic demography (Prugnolle and De Meus 2002). A mitochondrial DNA assay was therefore developed and evaluated during the project.

2.1.1 Development of homologous microsatellite markers for gray triggerfish

2.1.1.1 Library preparation and screening

Microsatellite enriched libraries were generated based on the protocol developed by Bloor et al. (2001). High quality genomic DNA was extracted from spleen tissue of a freshly-caught gray triggerfish specimen using the DNA-easy extraction kit (Qiagen). Approximately 10 mg of the obtained DNA was digested with the restriction enzymes *HaeIII*, *AluI* and *RsaI* (Promega) respectively according to protocols recommended by the manufacturer. Digestion products were run on a 2% agarose gel and DNA fragments ranging in size between 400 bp and 1,000 bp were excised from the gel and purified from the gel matrix using the Qiagen gel extraction kit (Qiagen). The digested and size selected DNAs from all three enzymatic digestions were pooled and ligated with the blunt-end adaptors Oligo A and Oligo B described in Bloor et al. (2001);

successful ligation of adapters was verified by Polymerase Chain Reaction (PCR) as described in this publication. The digested and ligated DNA was hybridized with the 3'-biotinylated probes [CA]₁₂ and [TAGA]₉ (the resulting libraries are further referred to as the CA and TAGA libraries respectively). Hybridization was carried out in a total volume of 100 µl containing 7 µl of digested ligated DNA, 50 pM biotinylated probe, and 30 µl 20X SSC (3M Sodium Chloride, 300 mM Sodium Citrate, pH 7). The hybridization protocol included 10 min at 95°C, followed by one hour at the probe specific annealing temperature (50°C for both the CA probe and the TAGA probe). Fragments containing microsatellites were captured using the streptavidin-coated magnetic beads (Dynabeads M-280, Life Technologies) according to directions from the manufacturer. The obtained enriched DNA was amplified through PCR. The reaction was run in a 60 µl volume containing 10 µl of DNA in solution, 25 pmol of Oligo A, 12 nmol of dNTPs, 90 nmol of MgCl₂, 3 U of TAQ polymerase (Promega), and 1X of buffer (Promega). The thermal cycle consisted in an initial denaturation at 95°C for 3 min, 5 cycles of 95°C for 30 s, 50°C for 30 s, 72°C for 45 s, 15 cycles of 95°C for 30 s, 50°C for 30 s, 72°C for 55 s and a final extension of 30 min at 72°C. Amplified DNA fragments were ligated in a pGEM®-T cloning vector (Promega) and transformed into *Escherichia coli* competent cells (JM-109 strain Promega) following protocols from the manufacturer. Positive clones were selected on X-Gal Ampicillin plates and hand-picked for sequencing. A total of 501 colonies with insert (377 for the CA library and 124 for the TAGA library) were obtained and stored in LB-Ampicillin culture medium. The inserts of 206 of the positive clones were amplified through PCR using M13 universal primers. PCR reactions were performed in a total volume of 31 µl containing 1 µl of template DNA solution, 12.4 pmol of each primer, 6.2 nmol of dNTPs, 46.5 nmol of MgCl₂, 1.55 U of TAQ polymerase (Promega), and 1X of buffer (Promega). The thermal cycle consisted in an initial denaturation at 95°C for 3 min, 35 cycles of 95°C for 30 s, annealing temperature (T_a) for 30 s, 72°C for 1 min 30 s, and a final extension of 15 min at 72°C; T_a was 58°C during the first seven cycles and 55°C for the remaining 28 cycles. PCR products were purified using the Exo-SAP-IT kit (GE Healthcare) and sequenced using the Big-dye terminator kit (Applied Biosystems) v. 3.3 according to directions from the manufacturer. Sequencing reaction products were run on a 3730XL automated sequencer (Applied Biosystem) at the High Throughput Genomic Unit of Seattle (WA). Sequence chromatograms were edited and screened for microsatellite arrays in Sequencher (Genecodes Corporation) v. 4.10.1. One hundred and twenty five of the sequenced clones (122 for the CA library and 3 for the TAGA library) contained microsatellites bearing more than 9 repeats. Forty six of the clones carried a perfect microsatellite array of at least 10 repeats (45 from the CA library and 1 from the TAGA library). Nineteen imperfect microsatellites (18 in the CA library and 1 in the TAGA library) were also identified.

2.1.1.2 Primer design and testing

Primers were designed in regions flanking microsatellite arrays using the software Primer3 (Rozen and Skaletsky 2000; available at <http://frodo.wi.mit.edu/primer3/>) v. 0.4.0 and PCR conditions were optimized for annealing temperature using DNA from 6 specimens of gray triggerfish. PCR reactions were performed in a total reaction volume of 5 µl containing 6-12 ng of genomic DNA, 2 pmol of each primer, 1 nmol of dNTPs, 7.5 nmol of MgCl₂, 0.25 U of TAQ polymerase (Promega), and 1X of buffer (Promega). PCR cycles consisted in an initial denaturation at 95°C for 5 min, 35 cycles of 95°C for 30 s, T_a for 30 s, 72°C for 45 s, and a final

extension of 15 min at 72°C. PCR products were initially evaluated for amplification success and polymorphism by visualization on a high resolution NuSieve-GTG® agarose gel (Lonza). Loci showing consistent amplification success and evidence of polymorphic PCR products were selected for further evaluation. The 5' end of one of the two primers of each selected primer pair was labeled using one of the fluorescent dyes 6-Fam, Hex or Ned and the obtained labeled primer was used in subsequent PCR amplifications. PCR reactions were carried out at a series of annealing temperatures ranging from 52°C to 62°C to determine optimal annealing conditions for each primer pair. The obtained PCR products were run on a ABI-377XXL sequencer (Applied Biosystems) according to instructions from the manufacturer and evaluated for polymorphism and interpretability of the obtained electropherograms. The size of the PCR products was determined with reference to a size standard prepared as described in DeWoody et al. (2004).

2.1.1.3 Characterization of microsatellite loci

Genotypes were generated using optimized conditions in 35 gray triggerfish specimens collected by trawling along the Louisiana coast during the 2008 and 2009 fall SEAMAP surveys conducted by the National Marine Fisheries Service. The number of alleles (A), and the observed (H_o) and expected heterozygosity (H_e) statistics were calculated using the software ARLEQUIN v. 3.01 (Excoffier et al. 2005). Conformance of genotypic proportions to Hardy-Weinberg (H-W) expectations was tested using exact tests in the software GENEPOP v. 4.2 (Raymond and Rousset 1995, Rousset 2008). Probability values were estimated using a Monte Carlo Markov Chain method with 5,000 dememorizations, 500 batches, and 5,000 iterations per batch. The possible occurrence of artifacts impacting scoring including null alleles, large allele dropout and stuttering was tested using Microchecker v. 2.2.3 (Van Oostehout et al. 2004).

2.1.1.4 Multiplex PCR

The microsatellites were evaluated for multiplex PCR amplification using the approach described in Renshaw et al. (2006). Loci were grouped based on compatibility for simultaneous amplification in the same PCR reaction (i.e. loci are compatible if their PCR products have different size or if they are labeled with a different fluorescent dye; loci are preferentially associated if they have similar annealing temperatures). Reaction cocktails were optimized to achieve amplification of a panel of a minimum of 15 loci in 3 to 4 reactions.

2.1.1.5 Testing of microsatellites in queen and ocean triggerfish

All the markers that could be optimized and validated for genetic studies of gray triggerfish were tested on queen and ocean triggerfish specimens. Fin tissue samples from ocean triggerfish ($n = 14$) were provided by Capt. Mike Newman and fin samples from queen triggerfish specimens were provided by Dr. David Olsen ($n = 32$). Evaluation was conducted in single marker PCR and began at the optimal annealing temperature for amplification in gray triggerfish but additional temperatures were also evaluated when initial PCR tests were not successful. Methods for amplification, electrophoresis, scoring, and statistical evaluation of loci followed the procedures described above for gray triggerfish.

2.1.2 Mitochondrial DNA assay

We evaluated sequence variation in a 694 base pair fragment of the ND4 sub-unit of the NADH dehydrogenase encoded by mitochondrial DNA. This coding gene was successfully used in previous population genetic studies in a variety of organisms including reef fishes in the Gulf of Mexico (Pruett et al. 2005, Gold et al. 2011, Saillant et al. 2012). PCR amplification and sequencing of PCR products employed the universal primers NAP-2 (Arevalo et al., 1994) and ND4LB (Bielawski and Gold 2002). Polymerase Chain Reactions were conducted in a 25 μ l volume solution containing 25-50 ng of genomic DNA, 10 pmol of each primer, 1.25 U of TAQ polymerase (Promega Inc., Madison, Wisconsin), 5 nmol of dNTPs, 37.5 nmol of MgCl₂, 1X buffer (Promega). Amplification by PCR consisted in an initial denaturation at 95°C for 3 min, 35 cycles of 95°C for 30 s, 55°C for 30 s and 72°C for 1 min and 30 s, and a final extension of 15 min at 72°C.

The PCR products were purified using the Exo-SAP-IT PCR clean-up kit (GE Healthcare, Piscataway, New Jersey) and sequenced using the Big-Dye Terminator v. 3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, California) following manufacturer's protocol. Sequencing reaction products were run on a ABI-3730XL capillary sequencer (Applied Biosystems) at the High-Throughput Genomics Center in Seattle or on an ABI-3130 capillary sequencer in our laboratory.

Sequences were edited in the software Sequencher v. 4.10.1 (Gene Codes Corporation, Ann Arbor, Michigan) and aligned in MEGA v. 4.1 (Tamura et al. 2007). A 617 base pair fragment could be reliably scored in all individuals examined and was used for data analysis. The assay was evaluated in a preliminary survey of genetic variation in U.S. waters that is presented in objective 2.

2.2 Approach objective 2. Assessment of genetic variation and stock structure in U.S. waters

This project section consisted in analyzing genetic variation among gray triggerfish geographic populations in U.S. waters using the molecular markers developed in objective 1.

2.2.1 Sampling

Gray triggerfish form a relatively continuous population along the U.S. coast but could not be sampled continuously along the 3,000 km studied range. Our initial sampling design therefore focused on 7 localities in the northern Gulf of Mexico and South Atlantic regions. The location of samples grouped in each locality is presented in Figure 1. Specimens were collected during the summer and fall of 2008 and 2009. Specimens from South Texas (TX), East Texas (ETX), Louisiana (LA) and Mississippi (MS) were collected by trawling in conjunction with the National Marine Fisheries SEAMAP surveys. Additional samples from Mississippi were provided by anglers. Samples from the vicinity of Panama City, Florida (WF) were collected at recreational fishing docks or using traps during fishery independent reef fish monitoring programs. Samples from Southwest Florida (SWF) were obtained in conjunction with the reef fish monitoring program and the SEAMAP surveys implemented by the Florida Fish and Wildlife Research Institute. Samples from East Florida (EF) were obtained via hook and line fishing with the help of Capt. W. Taylor. Finally samples from offshore South Carolina (SC)

waters were collected by trapping in conjunction with the MARMAP survey implemented by the South Carolina Department of Natural Resources.

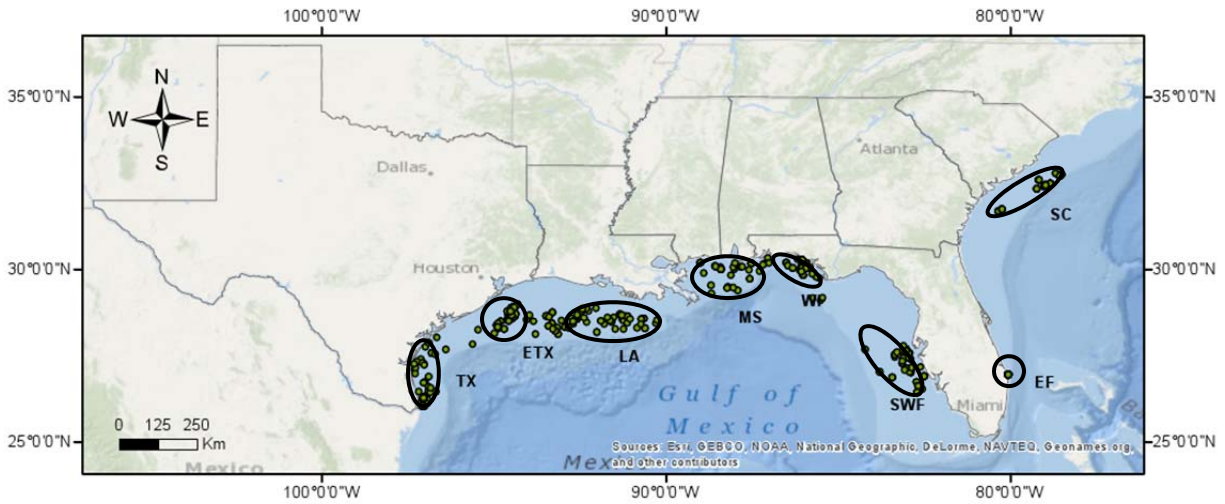


Figure 1. Sampling localities for gray triggerfish. TX: South Texas, ETX: East Texas, LA: Louisiana, MS: Mississippi, WF: West Florida, SWF: Southwest Florida, EF: East Florida, SC: South Carolina. Green circles represent individual sampling stations within a locality. Large black circles indicate initial grouping of samples. Additional SEAMAP samples are located outside circles delimiting initial localities.

Sample sizes ranged between 61 and 80 gray triggerfish per locality (Table 1). Tissue samples (finclips and/or muscle tissue) were collected on fresh or frozen specimens and preserved at room temperature in 95% ethanol or in a DMSO salt-saturated fixative (0.5 M EDTA, 20% Dimethyl sulfoxide, NaCl, ddH₂O) until DNA isolation except for the samples from South Carolina that were preserved in a sarkosyl urea lysis buffer (1% n-lauroylsarcosinate, 20 mM NaPO₄, 8 M urea, 1 mM EDTA). Additional samples from the northcentral and northwestern Gulf of Mexico obtained during the SEAMAP surveys were located between the localities targeted initially (Figure 1). These samples were analyzed and included in the analysis of isolation-by-distance (see results) in order to better reflect the continuous distribution of gray triggerfish along the coastline during estimation of the parameters of the model.

Table 1. Numbers of gray triggerfish specimens analyzed for each geographic region

Location	# samples
South Texas	72
East Texas	78
Louisiana	74
Mississippi	61
West Florida	77
Southwest Florida	77
East Florida	80
South Carolina	78
Additional SEAMAP samples (see Figure 1)	68

2.2.2 Assay

DNA was extracted from samples using a standard phenol-chloroform method (Sambrook et al. 1989) and all specimens were assayed at the 17 microsatellites developed in objective 1 and using methods described in this section. A pilot survey of the mitochondrial DNA fragment was conducted using samples from South Texas, Louisiana, West Florida, East Florida and South Carolina. Sample sizes for the mitochondrial DNA survey ranged between 27 and 32 gray triggerfish per locality (Appendix Table 4).

2.2.3 Data analysis

2.2.3.1 Genetic diversity

Microsatellites

Summary statistics at each microsatellites including number of alleles, allelic richness, and unbiased gene diversity were generated in Fstat v. 2.9.3. (Goudet 1995). Allelic richness is a measure of the number of alleles corrected for sample sizes and allows comparing samples with different sample sizes (El Mousadik and Petit 1996). Gene diversity was estimated as Nei's heterozygosity (Nei 1987). Allelic richness and gene diversity were compared among samples using Friedman's rank test as implemented in SPSS v. 20 software (IBM Corp.).

Departure from Hardy-Weinberg (H-W) equilibrium expectations in each locality sample was estimated as Weir and Cockerham (1984) inbreeding coefficient f . The probability that f differed from zero was determined using exact tests in GENEPOP v. 4.2 (Raymond & Rousset 1995; Rousset 2008). The exact probability values were estimated using a Monte Carlo approach based on 10,000 dememorizations, 500 batches, and 5,000 iterations per batch. The occurrence of scoring errors due to null alleles, stuttering or large allele dropout was evaluated in Microchecker v. 2.2.3 (Van Oosterhout et al. 2004).

Mitochondrial DNA

A statistical parsimony network of mtDNA haplotypes was generated in TCS v. 1.21 (Clement et al. 2000). Summary statistics including number of haplotypes, haplotype diversity and nucleotide diversity were computed in DNAsp v. 5.10 (Rozas et al. 2003). Homogeneity among samples in number of mtDNA haplotypes and haplotype diversity was tested using a bootstrap resampling approach (Karlsson et al. 2008). Under the null hypothesis of homogeneity, the number of different haplotypes or haplotype diversity statistics observed in one locality sample does not differ significantly from that in a random sample of the same size from the overall dataset (all localities pooled). POP TOOLS (a free-add in software for EXCEL, available at <http://www.cse.csiro.au/poptools/index.htm>) was used to generate random samples of 27 to 32 haplotypes from the overall dataset. Random sampling was performed 10,000 times and the average number of observed haplotypes/haplotype diversity statistics and their upper (0.975) and lower (0.025) percentiles recorded. Observed statistics in individual localities were significantly different from expected under the null hypothesis of homogeneity if they lay outside the bounds of the 95% bootstrap confidence interval (CI) obtained.

2.2.3.2 Analysis of spatial genetic variation

Microsatellites

Divergence among samples was estimated as Weir and Cockerham (1984) θ as computed in Genepop. The probability that θ differed from zero was determined using exact homogeneity tests employing the same Monte Carlo parameters as above for Hardy-Weinberg tests. Sequential Bonferroni correction (Rice 1989) was applied to account for multiple tests performed simultaneously.

Spatial patterns of population structure were examined via pairwise exact tests of population differentiation in Genepop. Population structure was also examined using Spatial Analysis of Molecular Variance (SAMOVA, Dupanloup et al. 2002) and Bayesian clustering in Structure (Pritchard et al. 2000). SAMOVA employs a simulated annealing algorithm to optimize allocation of N geographic populations into K groups ($2 < K < N$) in order to maximize the proportion of total genetic variance due to genetic variation among the inferred groups. A total of 100 simulated annealing processes were performed using the software SAMOVA v. 1.0 available at <http://web.unife.it/progetti/genetica/Isabelle/samova.html> in order to determine optimal allocation of the 8 geographic samples into up to 7 groups. Bayesian clustering was run using the correlated allele frequencies model (Falush et al. 2003). Chains were run for a total of 10^7 steps following 10^6 steps discarded as burn-in. Multiple values of K were tested and compared on the basis of the obtained posterior probabilities as described in Pritchard et al. (2010).

Considering the continuous distribution of gray triggerfish, we examined structuring according to an isolation-by-distance model. Both spatial autocorrelation analysis (Smouse and Peakall 1999) and the method developed by Rousset (2000) and Leblois et al. (2004) were employed and revealed weak structuring consistent with isolation-by-distance. Only results of the second approach are presented here as this method also allows estimating dispersal parameters based on existing theory of isolation-by-distance (Rousset 1997). The genetic distance between pairs of individuals was estimated as the \hat{e} statistics (Watts et al. 2007) computed in the software Genepop v. 4.2 (Raymond & Rousset 1995, Rousset 2008). The \hat{e} statistics is more powerful in cases where the spatial pattern of population structure is weak (Watts et al. 2007) as is the case in the present study (see findings section). Considering the shelf habitat used by gray triggerfish, dispersal and individual coordinates were considered on a one dimensional axis (along the coastline and approximated using mid-shelf transects) for computations of geographic distances between individuals and isolation-by-distance statistics. The analysis of isolation-by-distance was conducted primarily using data obtained on specimens ($n = 430$) collected between South Texas and West Florida because this portion of our sampling design approached best a continuous sampling along the coastline as recommended to infer parameters of the model (Leblois et al. 2004). Because estimation of the parameters of the isolation-by-distance model is biased when the geographic distance between samples being compared is greater than $0.56\sigma/\sqrt{2\mu}$ where σ is the standard deviation of parental position relative to offspring position and μ is the mutation rate (Rousset 1997), a bootstrap resampling approach was used to investigate the effect of the spatial scale of sampling on estimates of the slope of the isolation-by-distance relationship and σ . Subsamples were drawn by resampling sets of 100 individuals located within subsections

of the lattice of various lengths and the slope of the relationship between genetic and geographic distance (b) was estimated for each resampled dataset.

The slope of the linear regression between geographic and genetic distances (b) was used to calculate σ , given the effective population density (D), using the relationship (Rousset 1997)

$$\sigma = \sqrt{\frac{1}{4Db}} \quad \text{equation 1}$$

Inferences on σ thus require information on population density. Two approaches were taken to obtain values for D and discuss values of σ and the distribution of dispersal distances. An upper bound of D is given by the census population density. The census density of gray triggerfish in U.S. waters was calculated based on average landing data in the Gulf of Mexico during the period 2005-2012 obtained from the recreational fisheries statistics database of the fisheries statistics division of the National Marine Fisheries Service and accounting for fishing mortality rates between 0.43 and 0.53 reported in the SEDAR assessment. The census number for the Gulf of Mexico was divided by the length of the coastline, approximated using mid-shelf transects, to derive an estimate of census density. Effective density was also directly estimated using genetic data. Considering the homogeneity in allele frequencies throughout U.S. waters (see below), an estimate of effective population size for the regional metapopulation was generated using Approximate Bayesian Computations in the software DIYABC v. 2.0 (Cormuet et al. 2008). The effective size estimate was applied to the entire section of coastline surveyed (3,100 km) to derive an estimate of effective density.

Because the genetic consequences of dispersal depend on the shape of the distribution of dispersal distance (Rousset 2008), a simulation approach, similar to that employed by Puebla et al. (2012), was taken to determine the parameters of dispersal distance distributions yielding isolation-by-distance slopes consistent with that estimated from the empirical dataset. Coalescent simulations were implemented in the software IBDSim v. 1.4 (Leblois et al. 2009) considering various distribution functions. Simulations were performed considering a one dimensional lattice of 10,000 km with absorbing boundaries and 1,500 nodes were sampled corresponding to the studied portion of the northern Gulf of Mexico where a semi-continuous sampling was available (South Texas to West Florida) as discussed above. Simulated datasets included 17 unlinked loci with a range in allele size of 4 to 50 determined based on the empirical dataset. A two-phase mutation model (TPM) at a mean mutation rate of 5×10^{-4} was considered with 90% of single-step mutations, and a geometric variance of multi-step mutations of 10. The simulated datasets were processed for isolation-by-distance analysis as described above. Parameters for each of the dispersal distribution functions (Geometric, Pareto, and Sichel) were adjusted to determine values leading to isolation-by-distance slopes b similar to those obtained with the empirical dataset.

Mitochondrial DNA

Genetic homogeneity among localities was tested using analysis of molecular variance (AMOVA, Excoffier et al. 1992) as implemented in ARLEQUIN v. 3.01 (Excoffier et al. 2005) using the number of pairwise differences between haplotypes as a molecular distance. Significance of the genetic variance among localities and Φ_{CT} was assessed using 10,000 permutations of haplotypes. Population subdivision was further evaluated using spatial analysis

of molecular variance (SAMOVA, Dupanloup et al. 2002). Parameters were the same as those used for analysis of the microsatellite dataset except that optimal allocation of the 5 geographic samples into up to 4 groups was considered.

Possible occurrence of restricted gene flow and isolation-by-distance was evaluated using spatial autocorrelation analysis (Smouse and Peakall 1999; Peakall et al. 2003) as implemented in GenAlEx v. 6.3 (Peakall and Smouse 2006). The spatial autocorrelation coefficient (r) was computed using a linear, one-dimensional coastline geographic distance and the number of pairwise differences between haplotypes as a genetic distance as above. With isolation by distance, the estimated value of r among proximal localities (small distance classes) is significantly greater than zero and decreases with increasing distance between localities. Because estimation of spatial autocorrelation is influenced by the size of distance classes, r was computed based on a series of increasing distances between pairs of localities following Peakall et al. (2003). Distance classes were determined so that at least two pairs of localities per distance class would be included in the computation of r . Significance of r was determined via 10,000 random permutations of haplotypes among localities and using a one-tailed test (i.e., the observed r is significantly greater than zero if it lies beyond the upper 95% bound of the distribution of r in randomized datasets).

2.2.3.3 Effective population size and demographic history

Effective population size was investigated in order to provide values to estimate effective density for the analysis of isolation-by-distance conducted above and also to provide a first evaluation of potential signatures of historical events such as demographic bottlenecks or population growth at the regional level. In the absence of significant spatial structuring, gray triggerfish was treated as one single population. As discussed later in this report, the high dispersal and connectivity observed in gray triggerfish suggests that gene flow from foreign populations in particular in the Caribbean region and South America may influence genetic variation and demography of U.S. stocks. Samples from those populations have not yet been obtained to date so the analysis conducted here will require updating accounting for populations possibly exchanging migrants with U.S. stocks in order to correct estimates.

Microsatellites

Approximate Bayesian Computation modelling of demographic history based on the coalescent theory was applied to generate an estimate of effective population size for gray triggerfish using DIYABC v. 2.0 (Cornuet et al. 2008). Scenarios considering a constant N_e through recent evolutionary time, or one or two historical changes in population sizes were considered. Uniform priors were set for all parameters (effective population sizes at various stages of the scenario, timing of the events) encompassing a broad range of possible values (prior N_e was 100 – 100,000 and up to 100,000 generations were allowed between change in population size events).

The models for each scenario were simulated in 10^6 datasets and compared to the observed data on the basis of four summary statistics (mean number of alleles, mean expected heterozygosity, mean allele size variance and mean ratio of number of alleles over the range in allele sizes). The posterior support for each model was estimated as described in Cornuet et al. (2008). A model checking analysis was performed by comparing the first three axes from a Principal Component Analysis on observed summary statistics to those obtained from 1,000 simulations based on the

posterior predictive distribution of the best-fitting model (Cornuet et al. 2010).

Mitochondrial DNA

Neutral evolution of the mtDNA locus was tested via Fu's (1997) *FS* statistic and Fu and Li's (1993) *D** and *F** statistics, as implemented in the DNA-SP package (Rozas et al. 2003). Significance of *FS*, *D** and *F** was assessed using 10,000 coalescent simulations in DNA-SP. Demographic history of gray triggerfish was examined using the Bayesian Skyline Plot (BSP) approach (Drummond et al. 2005). The BSP is a method for estimating past population dynamics (i.e., change in effective population size N_e) through time from a sample of molecular sequences without dependence on a pre-specified parametric model of demographic history. The method allows discovering novel demographic signatures that are not readily described by simple demographic models. Bayesian Skyline Plots of gray triggerfish effective population size through time were constructed using a Monte Carlo Markov Chain (MCMC) sampling algorithm, as implemented in BEAST v. 1.4.8 (Drummond and Rambaut 2007) available at <http://beast.bio.ed.ac.uk/>. The underlying population size function of the BSP was fitted using a piecewise linear function of population size change based on 10 control points in BEAST. Gene trees were inferred using a general time-reversible substitution model with Gamma distributed variation in substitution rate among sites and a proportion of invariant sites. Substitution rates were estimated allowing for rate variation along the branches of the gene tree (relaxed molecular clock uncorrelated setting) with rates in each branch drawn from a log normal distribution. Three MCMC chains were run each involving 1.6×10^8 steps. The software Tracer v. 1.4.1 available at <http://tree.bio.ed.ac.uk/software/tracer/> was used to visualize trace files generated in BEAST and construct the skyline plot. Examination of the time series of the MCMC plot, the effective sample sizes for each parameter computed in Tracer, and the overall consistency of the three MCMC chains indicated that runs had converged to the posterior distribution. The first $2-3 \times 10^7$ steps were discarded in order to avoid influence of starting values on the estimated posterior distributions. The three runs gave consistent results and were combined to derive final posterior distributions and BSP for the dataset. Assignment of a time scale to the population size estimates was achieved by calibrating the rate of molecular evolution. A range of substitution rates of 1-1.5% per Million Years (MY) was considered based on molecular-clock calibrations of the mitochondrial cytochrome c oxidase (1.2%/MY) and NADH-2 and ATPase6 (1.3%/MY) genes developed for several geminate species pairs of fishes (Bermingham et al. 1997). Computations were performed at the Computational Biology Service Unit at Cornell University.

2.3 Approach objective 3. Initiate study of the connectivity of U.S. stocks with foreign stocks in the East Atlantic

Considering the high dispersal potential of gray triggerfish and the apparent homogeneity within the sampled range in U.S. waters, an additional important question that was raised during the course of the project is whether U.S. stocks are connected to foreign populations and partially dependent demographically on these stocks. A first axis of connectivity of importance to management of U.S. stocks is through the Caribbean Current. The Caribbean Current originates from the North Brazil and Guyana Currents and flows to the West entering the Gulf of Mexico as the Loop Current, potentially connecting the northern part of South America and the Caribbean

to U.S. stocks. An alternative axis of connectivity in that region is through the Antilles Current. Surface current data obtained from the US Coast Guard's Mariano Global Surface Velocity Analysis (MGSVA) 1.0 (Mariano et al. 1995) suggest that transfer of larvae from Northern Brazil to U.S. waters is possible within the 4 to 7 months of pelagic dispersal. Connectivity through the Antilles is also possible although currents are slower and direct transfer within one generation is less likely.

U.S. stocks could also be connected to Europe via transport through the Gulf Stream. A third hypothesis is that U.S. stocks are partially connected to populations in West Africa through the Equatorial Currents. Sampling during the project aimed to begin addressing these hypotheses.

2.3.1 Sampling

Of particular interest was the documentation of populations from the Caribbean region and South America as these populations are a likely direct source of migrants to U.S. stocks as discussed above. Unfortunately, gray triggerfish are infrequent in the Caribbean and efforts by collaborators to collect them have not been successful so far. We were however able to analyze fin clip samples from East Atlantic populations (Southwest of France $n = 64$ and Benin $n = 72$) and conducted a first examination of connectivity with these populations.

2.3.2 Assays

All samples were assayed at the 17 microsatellites developed in objective 1 and a subsample of 37 (France) and 39 (Benin) was assayed at mitochondrial DNA. Methods followed protocols described in objective 1.

2.3.3 Data analysis

Data analysis followed the methods described in objective 2 for summary statistics, analysis of genetic diversity and tests of Hardy-Weinberg equilibrium (microsatellite data) and description of diversity statistics and phylogenetic relationships of haplotypes in a TCS network (mtDNA).

The two populations (France and Benin) were compared to U.S. populations treated as one group based on results of analyses conducted in objective 2. Weir and Cockerham (1984) θ was computed overall populations and between pairs of populations and the significance of θ estimates was tested using exact test in Genepop using parameters reported in objective 2.

The software Structure (Pritchard et al. 2000) was used to examine occurrence of structuring within the dataset. This approach uses multilocus genotypes to detect distinct populations and optimizes assignment of individuals to populations in order to achieve minimal departure from Hardy-Weinberg equilibrium and linkage disequilibrium within the inferred clusters. The model accounting for correlated allele frequencies was used in computations (Falush et al. 2003).

Chains were run for a total of 10^7 steps following 10^6 steps discarded as burn-in. Multiple values of the number of clusters (K) were tested and compared on the basis of the obtained posterior probabilities as described in Pritchard et al. (2010).

Spatial analysis of mtDNA variation utilized Analysis of Molecular Variance as described in objective 2.

Historical demography was examined in the three populations accounting for migration using the Bayesian coalescent approach implemented in the program MIGRATE (Beerli and Felsenstein 1999, 2001; v. 3.6, <http://popgen.scs.fsu.edu/Migrate-n.html>). The parameter theta (θ) and M were estimated where θ is defined as $4N_e\mu$ for microsatellites and $N_e\mu$ for mtDNA (N_e is effective population size, μ is the per-locus mutation rate) and M is the average mutation-scaled migration rate. θ and M estimated by this method represent averages integrated over the time to the common ancestor of all sampled alleles. The full effective size and migration matrix of θ and M was estimated in a Bayesian framework using uniform priors for all estimated parameters allowing for a broad range of values for both θ (0-1.2 for mtDNA, 0-100 for microsatellites) and M (0- 20,000). Markov Chain-Monte Carlo (MCMC) coalescent simulations of gene-trees were used to generate posterior distributions of parameters. Parameter estimates were derived based on the results of three replicate runs. Each replicate run employed 4 long chains (1.5×10^7 trees sampled) with an adaptive heating scheme and discarding the first 10^4 trees as burn-in. Runs with the microsatellite dataset employed a subsample of 30 individuals and assumed a Brownian mutation model allowing for variation in mutation rate among loci. Run with mtDNA assumed a Felsenstein (1984) model with a transition:transversion ratio of 8.

3. Findings

3.1 Markers development

3.1.1 Microsatellites

3.1.1.1 Development of PCR primers and testing

The characteristics of the 21 primer pairs that achieved successful amplification of a scorable PCR product are presented in Appendix 1. The number of alleles detected (A) averaged 14.3 and ranged between 4 (locus BC16) and 27 (locus BC46) and the expected heterozygosity averaged 0.795 and ranged between 0.502 (locus BC49) and 0.967 (locus BC46). Genotypic proportions did not depart from H-W equilibrium expectations except for loci BC20 and BC36. Analyses in Microchecker indicated possible occurrence of null alleles and/or stuttering at locus BC20. No evidence for artifact impacting scoring at locus BC36 was detected during analyses in Microchecker. Locus BC9 displayed alleles differing by one base pair only, thus departing from the pattern of variation in number of repeats expected for microsatellites. This locus may therefore be of moderate usefulness for population genetic studies.

3.1.1.2 Multiplexing

On the basis of the above results, all loci except BC20 and BC9 were evaluated for multiplexing. Multiplex optimization resulted in 4 multiplex PCR assays for 17 microsatellites. The composition of the obtained multiplex panels is presented in Appendix 2. Locus BC16 could not be incorporated in multiplex PCR panels but could be amplified in single locus PCR and loaded for electrophoresis mixed with multiplex 40. This locus was therefore kept for the study. Locus BC36 was initially incorporated in the panel as part of multiplex 40 but subsequently removed as departure from Hardy-Weinberg and occurrence of scoring artifacts were detected in multiple

populations.

3.1.1.3 Evaluation of the developed microsatellites in queen and ocean triggerfish

All 17 microsatellites were tested in ocean and queen triggerfish. Initial amplification tests in queen triggerfish only led to 8 successfully amplifying loci but re-optimization of annealing temperatures yielded a total of 14 usable microsatellites (Appendix 3a). Only 4 of the microsatellites appeared suitable for population genetic studies of ocean triggerfish (Appendix 3b).

3.1.2 Mitochondrial DNA assays

Amplification of the ND4 fragment of mitochondrial DNA using the universal primers ND4LB and NAP2 was successful in all three species. Sequencing of PCR products yielded 617 bp of scorable sequence across U.S. samples and 599 bp across all assayed samples. In queen and ocean triggerfish, 620 bp and 629 bp of sequence could be scored unambiguously respectively. Further evaluation of the mtDNA assay proceeded for gray triggerfish during the pilot survey of population structure and is reported as part of objective 2 (section 3.2 below).

3.2 Analysis of stock structure in US waters

3.2.1 Mitochondrial DNA

Thirty eight unique mtDNA haplotypes (Appendix 4) were found among the 150 sequences obtained for U.S. populations. There were 43 base substitutions among the 38 haplotypes of which 32 were synonymous and 11 were non-synonymous changes with 42 transitions and 1 transversion. The statistical parsimony network revealed a star-like phylogeny of haplotypes that included three closely related haplotypes representing 59% of the samples and 23 singleton haplotypes (Figure 2). Haplotype diversity and nucleotide diversity were similar

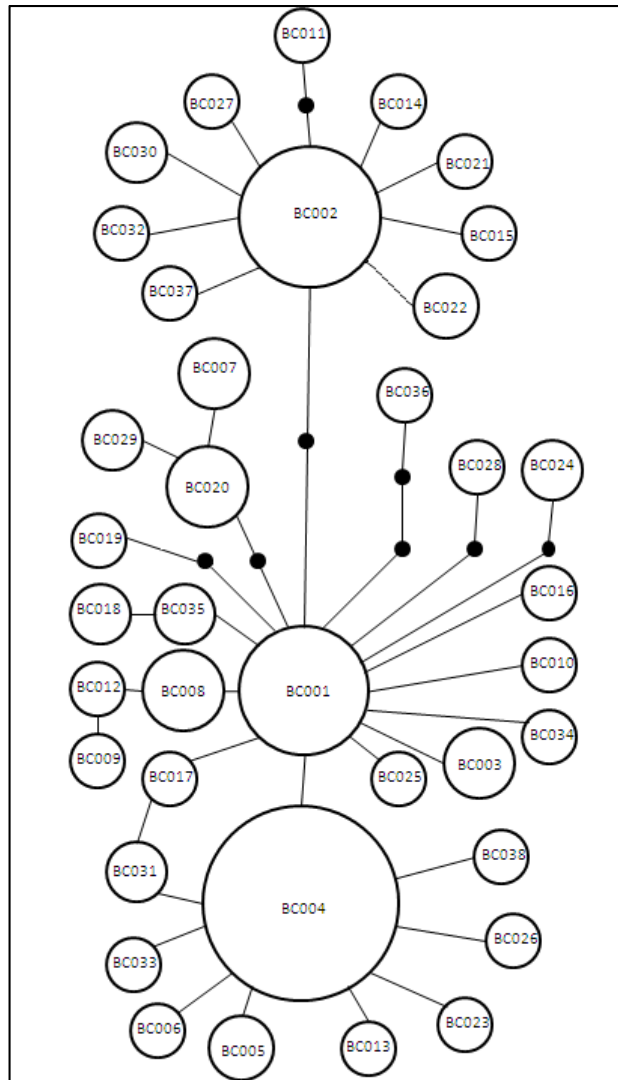


Figure 2. Statistical parsimony network of haplotypes detected in gray triggerfish. Numbers surrounded by circles correspond to individual haplotypes (detailed in appendix 2); the size of each circle reflects the frequency of occurrence of the corresponding haplotype. Solid lines between haplotypes, regardless of their length, correspond to single base-pair substitutions and dark circles represent unsampled intermediate haplotypes in the cladogram.

among localities and ranged between 0.815-0.907 for haplotype diversity and 0.00345-0.00451 for nucleotide diversity (Appendix 4). The number of haplotypes and haplotype diversity observed in each individual locality were all included in the 95% CI obtained in 10,000 bootstrap samples indicating that observed values did not differ significantly from those expected in random subsamples of the overall dataset.

Analysis of molecular variance did not reveal occurrence of significant heterogeneity of haplotype distributions among localities ($\Phi_{CT} = -0.0091$, $P = 0.805$). The among groups of populations component of molecular variance was maximized in SAMOVA when four groups were considered; the sample from Louisiana and Texas were associated in one group and the other three localities were inferred to constitute distinct groups (West Florida, East Florida and South Carolina respectively). The estimated molecular variance component among-groups was however negative (-0.08% of the total molecular variance) and did not differ significantly from zero ($P = 0.404$). Spatial autocorrelation coefficients did not differ significantly from zero for any of the distance class considered (data not shown).

3.2.2 Microsatellites

Summary statistics at each microsatellite locus in each population sampled are presented in Appendix 5. Nine out of 136 tests of H-W equilibrium gave significant outcome before Bonferroni correction but none of the tests remained significant after correction. In addition there was no evidence of artifacts impacting scoring such as null alleles, stuttering bands, and large allele dropout during analyses in Microchecker. All 17 microsatellites were therefore kept for further analysis.

The number of alleles (A) per locus ranged between 9 and 46 with an average of 25.3 and gene diversity was between 0.384 and 0.964 (average 0.783). Allelic richness and gene diversity did not differ significantly among localities ($Z = 4.9$, $df = 7$, $P = 0.678$; $Z = 7.2$, $df = 7$, $P = 0.405$ respectively). Exact homogeneity tests revealed significant differences in allele frequencies among localities overall loci corresponding to a very weak estimated F_{ST} value (0.0005). Only two tests at individual loci were significant before Bonferroni correction and none was significant after correction.

Pairwise comparisons confirmed very low levels of divergence among geographic populations and did not reveal a clear geographic pattern (Table 2).

Table 2. Estimates of pairwise F_{ST} (above diagonal) and probability that F_{ST} is different from zero from exact homogeneity tests (below diagonal). Bold font denotes significant P-values following Bonferroni correction.

	<i>TX</i>	<i>ETX</i>	<i>LA</i>	<i>MS</i>	<i>WF</i>	<i>SWF</i>	<i>EF</i>	<i>SC</i>
TX		0	0,0017	0,0005	0	0,0004	0,0019	0
ETX	0,0339		0,0005	0	0	0,0007	0,0013	0
LA	0,1204	0,0851		0,0013	0,0007	0,0021	0,001	0
MS	0,3774	0,0214	0,0342		0	0	0,0018	0
WF	0,5361	0,3378	0,2542	0,3784		0,0006	0,0005	0
SWF	0,1533	0,0009	0,0797	0,0706	0,1427		0,0019	0,0003
EF	0,2032	0,0988	0,0254	0,0097	0,6135	0,0027		0,0015
SC	0,5235	0,6767	0,9326	0,2266	0,6782	0,2823	0,1117	

TX: South Texas, ETX: East Texas, LA: Louisiana, MS: Mississippi, WF: West Florida, SWF: Southwest Florida, EF: East Florida, SC: South Carolina.

Spatial analysis of molecular variance did not reveal grouping of geographic samples in divergent units within the sampling surface and Bayesian clustering in Structure also failed to identify cryptic structuring within the dataset (data not shown).

Estimates of the slope of the isolation-by-distance model using subsets of the data encompassing increasing distance ranges revealed a high variance among the slopes estimated for resampled datasets at short spatial scale of observations. The mean and variance of slopes from resampled datasets stabilized between 3.4×10^{-8} and 4.4×10^{-7} when the sampled range was between 1,400 and 1,700 km (Figure 3).

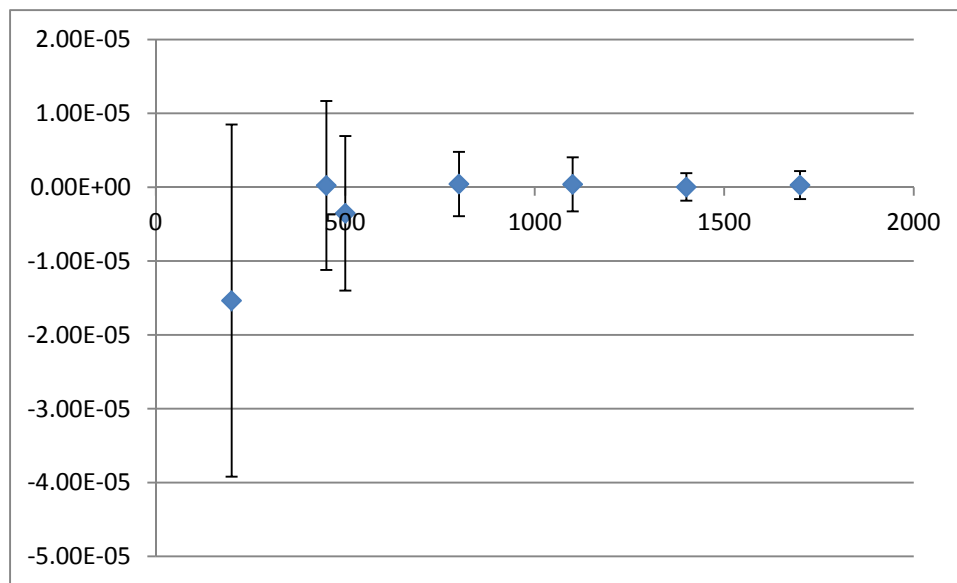


Figure 3. Average (\pm standard deviation) of the point estimate of the slope of the isolation-by-distance model in replicate subsamples of the dataset ($n = 100$) at increasing spatial scale of observation.

Accordingly, estimates were generated based on all available data for the area between South Texas and West Florida where the distribution of sample locations best reflected a continuous distribution. The obtained estimate was 3.1×10^{-8} (lower and upper bounds of the slope - 5.24×10^{-7} and 4.61×10^{-7}). Point estimates generated using greater portions of the dataset (i.e. including some or all remaining localities up to South Carolina) were all included within the bounds of the confidence interval described above.

The genetic estimate of D_e was generated based on estimates of current N_1 for the region (see section on demographic history below) applied to the entire lattice (from South Texas to South Carolina) yielding a value of 2.77 (1.46 – 17.83). The census density for the section of the Gulf of Mexico used to generate estimates of the isolation-by-distance slope was 175 giving a ratio effective to census density of 63.2 (9.8 – 119.9). Considering the genetic estimate of effective

density, the corresponding values of σ derived using equation 1 was 1628 (425, $+\infty$).

Because the genetic effects of dispersal depend largely on the shape of the distribution of dispersal distances, simulations were conducted to generate dispersal distributions, based on various families of distribution functions, yielding isolation-by-distance slopes similar to that obtained with the empirical dataset. The obtained dispersal distributions functions, their parameters, and the slope of the associated isolation-by-distance relationships are presented in Table 3. All compatible distributions involved mean dispersal distances greater than 238 km. Examination of cumulated distributions indicates that, 10% of dispersal events occurred at distance greater than 500 km for most distributions generated and a substantial fraction of simulated dispersal events occurred at large distance greater than 1,000 km (Appendix 6).

Table 3. Parameters of simulated distributions yielding isolation-by-distance slopes comparable to that of the empirical dataset (point estimate and upper bound).

Model	μ	σ (sim. /est.)	Slope	95%-	95%+	<i>P</i> -value
<i>D</i> = 3/km						
Pareto (M=0.99 ; n=1.02)	934	2151 / 441	4.28E-07	1.52E-07	7.69E-07	0.000
Pareto (M=0.96 ; n=0.602)	2789	4032 / 1667	3.00E-08	-9.28E-08	1.83E-07	0.103
Geometric (m=0.95 ; g=0.996)	238	344 / 465	3.85E-07	1.79E-07	5.91E-07	0.000
Geometric (m=0.97 ; g=1)	4850	5686 / 1145	6.36E-08	-1.20E-07	3.99E-07	0.322
Sichel (γ =-0.001 ; ξ =10000 ; Ω =0.002)	264	558 / 433	4.44E-07	2.71E-07	6.86E-07	0.000
<i>D</i> = 175/km						
Pareto (M=0.99 ; n=1.02)	934	2151 / 244	2.39E-08	-7.24E-08	1.30E-07	0.024
Geometric (m=0.95 ; g=0.996)	238	344 / 237	2.55E-08	-7.99E-08	1.30E-07	0.010
Sichel (γ =-0.001 ; ξ =10000 ; Ω =0.001)	284	600 / 301	1.57E-08	-1.13E-07	1.55E-07	0.118

D: population density, sim.: simulated, est.: estimated. μ : mean (simulated) dispersal distance, σ : standard deviation of parental position relative to offspring position

3.3 Demographic history

Based on the results of spatial analyses reported above, data from all U.S. localities were pooled for demographic analyses.

3.3.1 Mitochondrial DNA

Fu's *FS* and Fu and Li's *D** and *F** statistics were all significant (*FS* = -33.157, *P* < 0.0001; *D** = -4.59, *P* < 0.02; *F** = -4.24, *P* < 0.02) indicating departure from expected patterns under neutral genetic variation and mutation-drift equilibrium.

The mean of the posterior distribution obtained during Bayesian skyline reconstruction suggested occurrence of an initial rapid increase in historical population size followed by a phase of slower population growth and a third recent growth phase (Figure 4). Considering mutation rates of 1.0% per MY and 1.5% per MY, the initial growth phase would have occurred between 67,000-100,000 (considering 1.0%-1.5% mutation rates respectively) and 33,000-50,000 years ago. The

third phase would have begun 6,700 to 10,000 years ago and would have corresponded to a slight increase in population growth rate in recent times. The 95% highest posterior density interval however also included support for a reduction in population size during the latter phase (Figure 4).

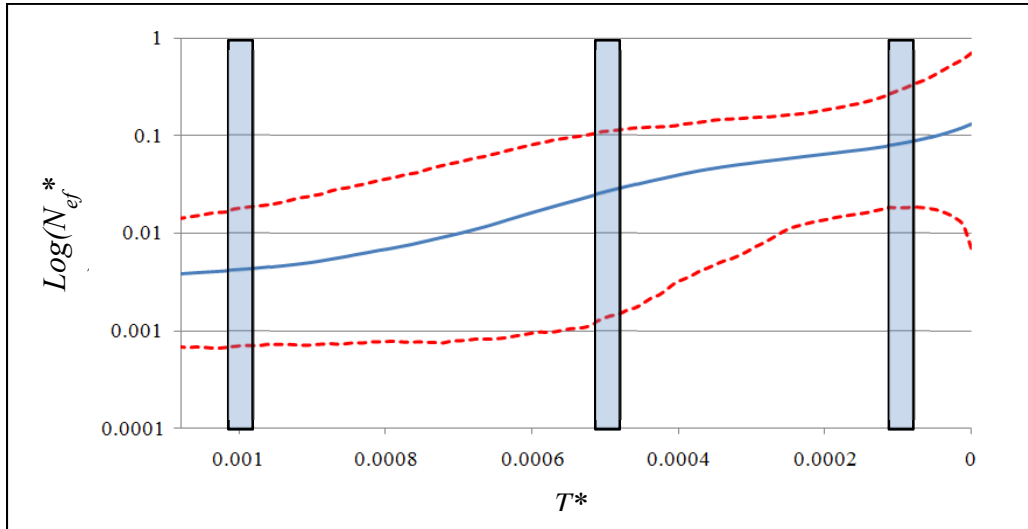


Figure 4. Bayesian skyline plot of decimal logarithm of gray triggerfish female effective population size (N_{ef}) through time (T , in number of generations) scaled by the substitution rate (μ in % per Million Year). The thick solid line represents the mean of the posterior distribution, and the dashed lines represent upper and lower limits of the 95% Highest Posterior Density interval. Vertical shaded strips mark historical changes in population growth rate discussed in text.

3.3.2 Microsatellites

All 3 models considered in the Bayesian coalescent approach in DIYABC received posterior support and, although scenarios 1 and 3 were marginally more supported, no scenario could be chosen with statistical confidence. A primary objective of the demographic analysis was to generate estimates of current effective population size that could be used to derive an estimate of population density. The summary distribution statistics for this parameter (N_1) based on the 3 scenarios are presented in Table 4.

Table 4. Summary of the posterior distribution of N_1 (current effective population size) obtained during coalescent-based Approximate Bayesian Computation analysis of demographic history

Parameter	Mode	95%-	95%+
Scenario 1	8.44×10^3	5.54×10^3	5.43×10^4
Scenario 2	1.81×10^4	1.05×10^4	7.19×10^4
Scenario 3	8.11×10^3	4.53×10^3	4.22×10^4

Because scenario 1 and 3 received highest posterior support and showed very similar posterior distributions for N_1 , the average mode under those two scenarios was used to generate a point estimate of population density.

3.4 Connectivity with foreign stocks

Genotype frequencies did not differ significantly from Hardy-Weinberg expectations in any of the three geographic regions following Bonferroni correction. Allelic richness and gene diversity did not differ significantly between the 3 regional samples ($0.07 < Z < 1.8$, $0.08 < P < 0.943$). Significant differences in allele frequencies were observed during exact tests of population differentiation among the 3 regions at 16 out of 17 loci after Bonferroni correction. Pairwise exact tests revealed that differences were essentially due to the Benin sample differing significantly from the other two ($0.16 < F_{ST} < 0.17$, $P < 0.0001$). The F_{ST} estimate between U.S. and France was negative (-0.0001) indicating similar distributions of alleles between the two regions. Bayesian clustering in Structure confirmed these results by grouping the U.S. and French samples into one clear cluster, the second cluster being essentially composed of Benin samples (Figure 5). There was no evidence for further structuring of the dataset.

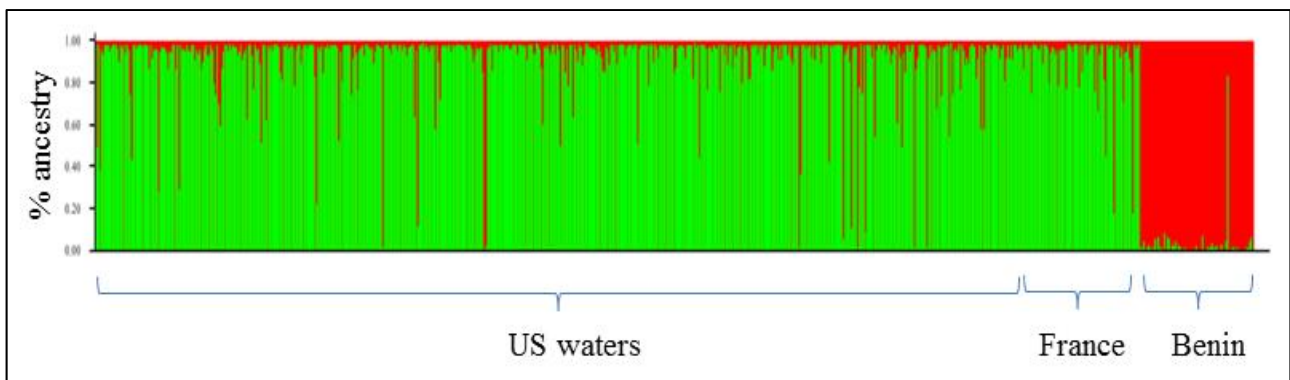


Figure 5. Bar chart illustrating the results of Bayesian clustering. Proportion of ancestry in the two inferred clusters in individuals sampled off U.S. waters, France and Benin

Analysis of the mitochondrial DNA dataset revealed clear differences in haplotype frequencies among populations ($\Phi_{ST} = 0.302$, $P < 0.0001$). Exact tests revealed no differences in haplotype frequencies between U.S. and France ($P = 0.774$) but highly significant differences in allele frequencies between Benin and both the U.S. and French localities ($P < 0.0001$), consistent with the pattern described above for microsatellites.

The sample from Benin showed significantly lower number of haplotypes, haplotype diversity and nucleotide diversity than samples from U.S. and France. All haplotypes from the Benin population were either BC020 or directly derived from this haplotypes (Figure 6). Haplotype BC020 was present in the French and U.S. populations but at a low frequencies.

Coalescent analyses in migrate revealed a higher effective population size in U.S. waters and France than in Benin at both classes of markers (Table 5). The difference was particularly pronounced at mtDNA where point estimates differed by almost two orders of magnitude. Estimates of M had very large confidence intervals and did not differ significantly among localities.

Table 5. Bayesian coalescent estimates of θ (mode and 95% highest posterior density intervals) for three geographic populations of gray triggerfish (United States, France and Benin) inferred from variation at mitochondrial DNA and microsatellite loci.

Mitochondrial DNA			
	θ ($N_e\mu$)	95%-	95%+
US	0.029	0	0.71
France	0.039	0.005	0.74
Benin	0.0004	0	0.006
Microsatellites			
	θ ($N_e\mu$)	95%-	95%+
US	6.7	4.2	9.1
France	4.4	2	6.7
Benin	0	1.4	2.9

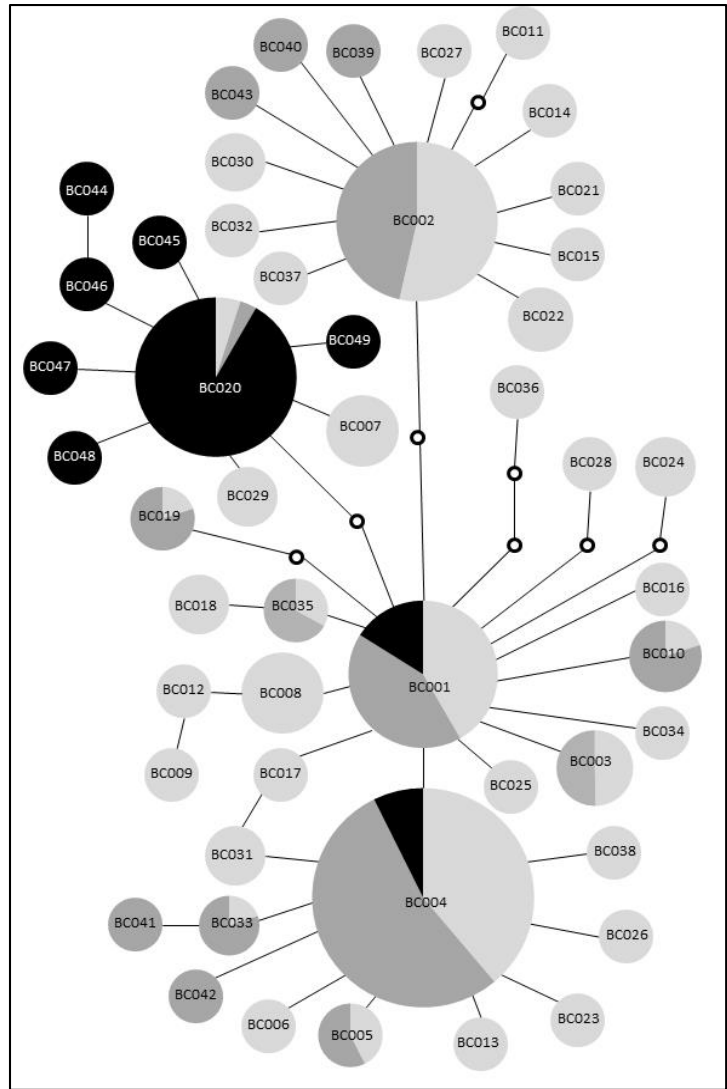


Figure 6. Statistical parsimony network of haplotypes detected in gray triggerfish. Numbers surrounded by circles correspond to individual haplotypes; the size of each circle reflects the average frequency of occurrence of the corresponding haplotype across populations (detailed in appendix 7). Solid lines between haplotypes, regardless of their length, correspond to single base-pair substitutions and open circles represent unsampled intermediate haplotypes in the cladogram. The relative abundance of each haplotype in U.S. waters, France and Benin is reflected by the proportion of each circle colored in light gray, dark gray and black respectively.

4. Interpretations and evaluation

4.1 Marker development

The first objective of this project was to develop a panel of molecular markers for genetic monitoring of gray triggerfish stocks that would include a minimum of 13-15 microsatellites and an assay of sequence variation at a mitochondrial coding gene. Twenty one new homologous microsatellites were developed and 17 of the loci among those found suitable for population genetic analysis could be optimized in 4 multiplex assays resulting in an efficient and powerful panel of markers for assessment of genetic variation and stock structure. The microsatellites were also tested in two other economically important triggerfish species, the queen and the ocean triggerfish. Fourteen of the microsatellites were found usable in *Balistes vetula*. The compatibility for multiplex assays of the 14 loci was not evaluated in this study and thus it is likely that a few additional markers will be needed in order to achieve a sufficient panel of multiplexed markers for population genetic surveys. Only four of the developed microsatellite markers were found usable in *Canthidermis sufflamen*. This species is in another genus and the lower transferability of gray triggerfish microsatellites is consistent with findings in other taxa where the success rate of cross amplification tends to decrease with increasing phylogenetic distance. Specific markers will therefore need to be developed in the future for genetic studies of ocean triggerfish.

The mitochondrial DNA assay could be developed using universal primers employed previously in other reef fishes of the Gulf of Mexico. This assay was used in a preliminary survey of genetic variation among gray triggerfish in the present work. Successful amplification and sequencing of the same fragment was achieved in a few test specimens of queen and ocean triggerfish suggesting that this assay could be successfully used in these two species as well.

4.2 Stock structure and connectivity

Analysis of spatial genetic variation revealed homogeneous frequencies of microsatellite alleles and mtDNA haplotypes across localities in U.S. waters. Additional spatial analyses including Spatial Analysis of Molecular Variance and Bayesian clustering provided no evidence of genetic discontinuity or cryptic stock structure within the same area. Altogether, these results suggest that no major barrier to gene flow is present across the sampling surface and that the stock in U.S. waters is not subdivided in discrete divergent units. This finding is tentatively consistent with the treatment of gray triggerfish as a single genetic stock across the sampled range encompassing the northern Gulf of Mexico and South Atlantic regions.

Tag and recapture studies indicate that gray triggerfish are highly sedentary as adults owing to territorial behavior (Ingram et al. 2001), suggesting that adult migration is not a likely mechanism promoting gene flow. In contrast, gene flow could be promoted through dispersal at the larval and juvenile stages. Although eggs are laid on the bottom in a nest, larvae and juveniles utilize the pelagic habitat and are found associated with seaweed and flotsam (mostly *Sargassum* spp.) (Aiken 1983) where they have been reported at sizes up to 3.15 in (Standard Length SL) and 5 in (Fork Length FL) by Wells and Rooker (2004) and Franks et al. (2007), respectively. More recently, a survey of age 0 gray triggerfish settling on artificial reefs in the northcentral Gulf of Mexico concluded that larvae and juveniles can spend 4 to 7 months in the

Sargassum beds (Simmons and Szedlmayer 2011).

While local recruitment could still be promoted if pelagic *Sargassum* aggregations remain near spawning habitats, dispersal over broad geographic distances is likely promoted over such long periods of transport. The extent and direction of dispersal would essentially depend on the direction and strength of oceanic currents encountered by the *Sargassum* during passive drift. As an example, *Sargassum* could be captured in the Gulf of Mexico and transported to waters off the U.S. East coast by the Loop Current (Ulanski 2008). Indeed, Gower et al. (2006) and Gower and King (2011) showed using satellite imagery that *Sargassum* originates in the Gulf of Mexico in spring and is advected in the Atlantic during the summer which corresponds to the period of gray triggerfish larval and juvenile transport. This pattern seemed consistent across years and would promote gene flow from the Gulf of Mexico to the South Atlantic and possibly further north through the Gulf Stream. Accordingly, a predominantly unidirectional pattern of gene flow (Gulf to Atlantic) would be expected. There was no evidence for such pattern in the genetic data with no detectable difference in allele frequencies and gene diversity indices between Gulf and Atlantic localities. However, considering the homogeneity observed across the sampling surface and the potential major role of long distance dispersal (see below), asymmetric gene flow along the Loop Current-Gulf Stream axis may occur yet might not be detectable with the data. Within the Gulf, while average current patterns could lead to some predictions regarding the directionality of gene flow in different parts of the Gulf (see e.g. seasonal averages at http://oceancurrents.rsmas.miami.edu/atlantic/loop-current_3.html), events such as storms and Hurricanes that are frequent during the summer and early fall, when gray triggerfish disperse, could easily offset the effects of predominant currents reflected in seasonal averages. Prediction of the overall direction and distance of pelagic transport is therefore a priori difficult considering that surface currents patterns that may be involved are diverse and likely vary in direction and strength during the course of the protracted spawning season (and subsequent pelagic phase) of gray triggerfish.

The finding of spatial genetic homogeneity of gray triggerfish from South Texas to South Carolina contrasts with patterns of population genetic structure reported in other marine fishes in the region. Divergence between Atlantic and Gulf populations of several species was reported by Avise et al. (1992) and was hypothesized to reflect a variety of factors including historical processes, discontinuous habitats, current patterns, life history, and behavior. More recently, significant structuring among geographic populations of gray snapper (*Lutjanus griseus*) and lane snapper (*Lutjanus synagris*) was reported by Gold et al. (2009) and Karlsson et al. (2009). Divergence involved the eastern versus western Gulf of Mexico (gray snapper and lane snapper) and the Gulf of Mexico versus the east coast of Florida (gray snapper). Recent and on-going studies of the red snapper (*Lutjanus campechanus*), while indicative of homogeneous distributions of genetic variants across large geographic scales in the Gulf, revealed weak structuring of Gulf *versus* Atlantic red snapper (Hollenbeck et al. Submitted) and a weak but significant isolation-by-distance signal in the Gulf of Mexico (Saillant et al. 2010). The three snapper species are sedentary at the adult stage as gray triggerfish are, and the red snapper occupies offshore habitats similar to those used by gray triggerfish. However, the snapper species differ from gray triggerfish in their dispersal potential during early life stages: gray, lane and red snapper larvae are pelagic but do not use the *Sargassum* as a habitat and also only disperse for a much shorter amount of time (the planctonic phase of the snapper larvae is estimated to last 20-40 days before settlement on benthic habitats occurs, Lindenman et al. 2000

versus 4 to 7 months for the gray triggerfish as discussed above). The homogeneity across the entire Gulf and South Atlantic regions, encompassing more than 3,000 km of coastline reported here thus suggests that the dispersal of juveniles by the *Sargassum* is effective at promoting gene flow and connectivity across extensive geographic ranges.

We do note that there are a number of caveats to the inference that gray triggerfish form a single stock in U.S. waters. As discussed by Pruett et al. (2005), a first issue is that spatial genetic homogeneity could stem from high gene flow within the area, but also from a non-equilibrium situation where regional populations would currently be partially isolated but there would not have been sufficient time since isolation to allow for accumulation of detectable genetic differences. Distinguishing between the two hypotheses is difficult as both may result in homogeneous present-day haplotype and genotype frequencies as is observed in gray triggerfish. Another caveat to the single stock hypothesis is that very limited level of gene flow (only a few migrants every generation) is sufficient to maintain genetic homogeneity (Waples 1998). Thus populations may be demographically independent while displaying no detectable genetic divergence. Finally, the microsatellite markers that we used are well suited to characterize gene flow and genetic drift but do not reflect local genetic adaptation. Deployment of genome scans based on large panels of SNPs would be useful to determine if local adaptation of geographic populations (e.g. Gulf *versus* South Atlantic) does occur and incorporate this information during assessment of genetic stock structure.

The second component of our analysis of spatial genetic variation aimed to investigate the spatial scale of dispersal using isolation-by-distance models based on individual data (Rousset et al. 2000, Leblois et al. 2004, Watts et al. 2007). The method was deployed in recent studies to infer dispersal parameters in other reef fishes where little divergence could be detected across large geographic areas (Puebla et al. 2012). Estimates of the slope of the isolation-by-distance model in gray triggerfish were consistent with large neighborhood sizes and the corresponding simulated dispersal distribution functions all involved substantial fractions of dispersal events over several hundreds of kilometers. A first issue associated with estimation of the isolation-by-distance slope is that of the spatial scale of sampling. If the sampled geographic range is too short (e.g. all samples are taken within the neighborhood), then no divergence is to be detected and the estimated regression parameters will reflect sampling artifacts, meaningless with regards to the true dispersal function. On another hand if the geographic distance between samples is too large (greater than $0.56\sigma\sqrt{1/2\mu}$, Rousset 1997), the slope of the isolation-by-distance model is underestimated leading to an over estimation of the parameter σ and bias during inferences on the distribution of dispersal distances. We took an approach similar to that of Puebla et al. (2012) and generated replicated subsamples of 100 individuals from the dataset by resampling at increasing geographic scale. There was a large variance of estimated slopes in replicated datasets generated at small spatial scales (lower than 500 km in our dataset), an observation that was also made by Puebla et al. (2012) and suggests that these datasets were inappropriate to estimate dispersal parameters. The mean of estimated slopes and its variance tended to stabilize when the spatial scale of sampling was between 1,400-1,700 km. Estimates were therefore generated using all available samples from South Texas to West Florida (~1,700 km range) where sampling best reflected the continuous distribution of gray triggerfish. The inferred values of σ suggest that the sampled range used to derive estimates is appropriate (i.e. does not exceed $0.56\sigma\sqrt{1/2\mu}$) unless extreme demographic scenarios are considered (i.e. effective density equals census density) and

the upper bound of the 95% Confidence Interval for the slope generated using Rousset's (2000) measure of genetic distance is used.

A major source of uncertainty when estimating dispersal parameters from the slope of the isolation by-distance model lies in the determination of the population effective density. The effective density (its square root) directly influences the parameter σ and in consequence the inferred distribution of dispersal distances. Our genetic estimate of effective density differed from the census density estimate by a factor of 63.2 leading to much larger dispersal parameters when using the genetic estimate. The genetic estimate of D_e was generated by using a coalescent approach to estimate N_e and applying this estimate to the entire sampled area to derive an estimate of D_e , thus making the assumption that our estimate reflects the size of a highly connected regional metapopulation from South Texas to South Carolina. Methods to derive an estimate of D_e from genetic data in other studies included estimates of contemporaneous effective population size (e.g. the temporal method in Pinski et al. 2010) or approximate Bayesian Computations similar to our approach in the study of Puebla et al. (2012). One difficulty when implementing this approach is to determine the spatial scale at which the genetic estimate of N_e applies. Gray triggerfish are relatively abundant throughout the surveyed area in U.S. waters but are very infrequent throughout the Caribbean, which led to apply the N_e estimate to the U.S. lattice. The exact extent of the gray triggerfish metapopulation that involves U.S. stocks would however need to be formally determined through an assessment of the occurrence of foreign populations, in particular in the Caribbean and South America, and their connectivity with U.S. stocks. Population density may also differ among regions and variation in density among regions may need to be accounted for as well.

Our second approach was to use census density (D) as proxy for D_e . D was approximated from catch data and its direct use in equation 1 makes the assumption that effective density equals census density. Demographic approaches are available to calculate effective population size based on census size and life history data (Nunney and Elam 1994) but this method requires knowledge of demographic parameters that are difficult to assess with confidence in marine species such as the gray triggerfish and these data are currently unavailable. Life history parameters lead to a reduction of D_e/D in almost all situations and thus the direct use of D corresponds to an extreme higher bound for D_e and is almost certainly too high. The value of the ratio D_e/D (or N_e/N) is not a straightforward question in marine species as some studies have reported effective size to be several orders of magnitude smaller than census size. The present point estimate for gray triggerfish corresponds to a D_e/D ratio of approximately $1.6 \cdot 10^{-2}$, a value intermediate between the extremely low ratios (10^{-3} to 10^{-5}) reported by Hauser et al. (2003) and Turner et al. (2002) in two marine fishes and the range (> 0.25) expected in most situations based on demographic approaches (Nunney and Elam 1994). Considering that direct information on life history parameters that include variance in survival will be very challenging to obtain in gray triggerfish, a better characterization of genetic variation in the West Atlantic (including southern populations) and information on the location and density of gray triggerfish populations in the Caribbean and South America would be beneficial to improve the genetic estimates of D_e .

Recent studies in other reef fishes with pelagic larvae reported that the spatial scale at which larval transport operates at an ecologically relevant magnitude is very limited with connectivity envelopes often less 100 km (Cowen et al. 2006, Puebla et al. 2012). However, these studies investigated species characterized by much shorter larval transport than the gray triggerfish (typically 20 to 40 days *versus* 4 to 7 months for gray triggerfish) and the study of Roberts

(1997) already pointed out longer dispersal distances when 2 month larval envelopes were derived, in particular if they involved transport through the Caribbean and Loop Currents. The combination of a much longer period of larval transport and opportunities for transport via fast currents such as the Loop Current could thus explain the much larger dispersal envelopes inferred for gray triggerfish in this work. Furthermore, the finding of homogeneous allele frequencies between the French and U.S. populations provide additional supports for the hypothesis that dispersal across large distances (e.g. from the U.S. coast to Western Europe) through pelagic transport is occurring in this species.

In consequence, further study encompassing a broader geographic sampling is needed in order to determine the spatial scale of this connectivity. Connectivity of U.S. ray triggerfish with populations in the Caribbean and South America could not be studied in the present work because samples have not been obtained to date but is a high priority for future studies.

The demographic history and status of the U.S. stock is difficult to discuss with confidence at this date. Analysis of the mtDNA dataset suggests occurrence of a historical expansion phase, likely following recent glacial retreats, a finding consistent with report in several other species. Estimation of contemporaneous effective population size was attempted in this work using the Linkage Disequilibrium Method (Waples 2006, Waples and Do 2010) but estimates were infinite or at least greater than 1,000, a range of values where the method is inaccurate. As discussed above, the high dispersal potential of gray triggerfish suggests that demographic dynamics is complex and will need to be evaluated accounting for foreign stocks if those are found to influence recruitment in U.S. waters. The clarification of the geographic scale of the metapopulation connected to U.S. stocks in the West Atlantic will be essential in order to account for connected stocks when estimating gene flow and effective population size and testing for potential reduction in effective population size (Storz and Beaumont 2002).

4.3 Connectivity with foreign stocks

Analysis of samples from the U.S. and France revealed very similar distributions of alleles and haplotypes between these two regions. This result supports the hypothesis that these two regions are connected via transport of juveniles through the Gulf Stream. On another hand, high levels of divergence was observed between the sample from Benin and both the U.S. and French samples suggesting that connectivity of the French-U.S. group with African populations is much more limited. Connectivity between West Africa and the U.S. would occur through the equatorial current which is very slow and one would expect several months of transport before juveniles can settle on western Atlantic habitat. The strong divergence at both classes of molecular markers with major shifts in allele/haplotype frequencies suggests that gene flow, if any, is very infrequent and limited. In particular, the major mtDNA clade that included almost all the Benin samples is very infrequent in the U.S. and French samples suggesting long term isolation and different evolutionary trajectories of the two groups. The Benin population also appears to display a smaller effective population size. This may be due to recent founder events or major historical Bottlenecks. A major decline in abundance of the species in West Africa has indeed been observed during the past few decades (Tore Stromme, Institute of Marine Research, Norway, Personal communication) and might contribute to the obtained estimates. However, the occurrence of several derived mtDNA haplotypes in Benin may suggest that the founder

event/Bottleneck is ancient and was followed by an expansion visualized in the star-like phylogeny of haplotypes in that region.

Although gray triggerfish are infrequent in Western Europe, the French sample displayed very large estimates of effective population size similar to those inferred for U.S. stocks. The inference of similar genetic diversity and effective size along with identical allele frequencies thus suggests that a large part of gray triggerfish found in Northern Europe might originate directly from U.S. and Caribbean stocks. The occurrence of breeding populations in Western Europe would therefore be worth investigating in the future to evaluate this hypothesis.

In summary, although the Gulf of Mexico and South Atlantic stocks appear to be currently managed and allocated separately, these two geographic stocks are likely at least partially dependent demographically. In particular, a potential seeding role of the Gulf of Mexico and Caribbean populations for the South Atlantic geographic stock would need to be further evaluated. Overall the degree of connectivity of U.S. stocks with populations in the Caribbean and South America will need to be investigated in priority in future studies.

Estimation of the slope of the isolation-by-distance model and dispersal parameters was limited in part by difficulties in determining the value of population effective density. Determination of the spatial scale of connectivity and estimation of the effective population size of the whole metapopulation directly connected to the U.S. stock, accounting for other populations partially connected to it, would be precious in order to improve estimates of effective density and dispersal parameters using genetic data.

A better mapping of gray triggerfish populations in regions of the Caribbean and South America where gray triggerfish may be present would also be helpful to build the model and account correctly for population density in the various sub regions.

5. Diffusion of the results and outreach

Two manuscripts were written during the course of the project and published in peer-reviewed journals. The first manuscript reports the results of the mitochondrial DNA survey in U.S. waters and the second reports the development of the homologous microsatellite markers. References for these two manuscripts are provided below.

A third manuscript reporting the microsatellite survey in U.S. waters is in progress and we anticipate submission to a peer-reviewed journal.

5.1 Peer-reviewed manuscripts

5.1.1 Published

- L. Antoni†, N. Emerick*, E. Saillant. Genetic variation of gray triggerfish (*Balistes caprisucus*) in US waters of the Gulf of Mexico and South Atlantic regions as inferred from mitochondrial DNA sequences. *North American Journal of Fisheries Management*, 31:714-721
- L. Antoni†, E. Saillant, 2012. Development and characterization of Microsatellite markers in the gray triggerfish (*Balistes caprisucus*). *Conservation Genetics Resources*, 4:629-631.

5.1.2 In preparation

L. Antoni†, E. Saillant. Connectivity over large geographic scales in an adult-sedentary reef fish with high potential for larval dispersal.

5.2 Presentations at conferences

5.2.1 Oral presentations

L. Antoni†, E. Saillant, 2012. Genetic variation of gray triggerfish in the US waters of the Gulf of Mexico and South Atlantic regions. Annual meeting of the American Fisheries Society, August 19-23 2012, Saint Paul, MN.

5.1.2 Poster presentations

N. Emerick*, E. Saillant, 2010. Testing for genetic isolation between Gulf of Mexico and U.S. Atlantic East Coast gray triggerfish using a mitochondrial DNA molecular marker. 63rd Annual meeting of the Gulf and Caribbean Fisheries Research Institute 2010, San Juan, Puerto Rico.

P. Carter*, A. Sward*, L. Antoni, E. Saillant 2010. Assessing genetic divergence between Gulf of Mexico and South Atlantic Gray triggerfish populations by Use of Microsatellite Markers. Poster presented at the Annual meeting of the Mississippi Academy of Sciences, Hattiesburg, MS, February 17-18 2011.

* Undergraduate student, † Graduate student

5.3 Student training

Four undergraduate students from the Mississippi Gulf Coast Community College (Nicholas Emerick, Avery Sward, Patrick Carter, and Megan Cox). Three of the interns (Emerick, Sward and Carter) were honor's students and presented poster presentations at meeting (one poster at a regional meeting and one at an international meeting).

One Ph.D student (Luca Antoni) is continuing his dissertation research on gray triggerfish genetics (completion anticipated in spring 2015).

5.4 Other outreach activities

A web page presenting genetic research on gray triggerfish is in development

A presentation to the public is scheduled for the coming annual workshop organized by the Center for Fisheries Research and Development at the Gulf Coast Research Laboratory (April 23rd). The workshop is attended by fishermen and aims to provide information on research findings as they apply to improve knowledge of fishes and fisheries.

During the course of the project the results obtained and the first manuscripts generated were communicated to the technical monitor and to the Southeast Data assessment and Review panel during the assessment update conducted in 2011. The present report and any coming manuscripts will be communicated to the coming SEDAR panel for South Atlantic gray triggerfish.

Acknowledgments

We thank B. Walling, D. Devries and B. Barnett (National Marine Fisheries Service, Panama City laboratory), K. Johnson, A. Debose and M. Hendon and the scientific staff of the National Marine Fisheries Service Mississippi Laboratory for assistance with collections in the northcentral and northwestern Gulf of Mexico, R. Chapman and M. Reichert for assistance with collections from South Carolina region, Capt. Bill Taylor and staff of the Black Dog Charters for assistance with collection of samples from Jupiter area, David Milly and Dominique Dirassar for collections in France, and Roger Djiman for assistance with collections in Benin. Jim Franks provided precious help with coordination of sampling. We also thank Capt. Mike Newman and Dr. Chris Koenig for collections of ocean triggerfish samples, Dr. David Olsen for collections of queen triggerfish specimens and Mrs. Jean Jovonovitch-Avillar for assistance with sequencing. Funding for this work was provided by the NOAA MARFIN project #NA09NMF4330150.

References

- Aiken K.A. (1983) The biology, ecology, and bionomics of the triggerfishes, Balistidae. Pages 191-205 in J. L. Munro, editor. Caribbean coral reef fishery resources. ICLARM Studies and Reviews 7.
- Allendorf F.W., Luikart G. (2007) Conservation and the genetics of populations. Blackwell Publishing Ltd., Oxford, UK.
- Arevalo E., Davis S.K., Sites J.W. (1994) Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosomes races of the *Sceloporus*. Syst. Biol. 43: 387-418.
- Avise J.C. (1992) Molecular population structure and the biogeographic history of a regional fauna: a case history with lessons for conservation biology. *Oikos* 63:62-76.
- Beaumont M.A. (2001) Conservation genetics. In: Balding DJ, Bishop M, Cannings C (eds) *The Handbook of Statistical Genetics*. John Wiley: New York, pp 779–812.
- Berli P., Felsenstein J. (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152: 763-773.
- Berli P., Felsenstein J. (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* 98:4563-4568.
- Bermingham E., McCafferty S.S., Martin A.P. (1997) Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. Pages 113-128 in T. D. Kocher and C. A. Stepien, editors. *Molecular Systematics of Fishes*. Academic Press, San Diego,

California.

- Bielawski J.P., Gold J.R. (2002) Mutation patterns of mitochondrial H- and L-strand DNA in closely related cyprinid fishes. *Genetics* 161: 1589-1597.
- Bloor P.A., Barker F.S., Watts P.C., Noyes H.A., Kemp S.J. (2001) Microsatellite Libraries by Enrichment. Protocol available at:
<http://www.genomics.liv.ac.uk/animal/MICROSAT.pdf>.
- Begg G.A., Friedland K.D., Pearce J.B. (1999) Stock identification and its role in stock assessment and Fisheries management: an overview. *Fish. Res.* 43:1-8.
- Carvalho G.R., Hauser L. (1995) Molecular genetics and the stock concept in fisheries. in Carvalho GR, Pitcher TJ (ed) *Molecular genetics in fisheries*. Chapman and Hall, London, pp55–80
- Clement M., Posada D., Crandall K. (2000) TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657-1660.
- Conover D.O., Arnott S.A., Walsh M.R., Munch S.B. (2005) Darwinian fishery science: lessons from the Atlantic silverside. *Can. J. Fish. Aquat. Sci.* 62: 730-737.
- Cornuet J.-M., Santos F., Beaumont M.A., Robert C.R., Marin J.-M., Balding D.J., Guillemaud T., Estoup A. (2008) Inferring population history with DIY ABC: a user-friendly approach to Approximate Bayesian Computation. *Bioinformatics* 24: 2713.
- Cornuet J.-M., Ravigne V., Estoup A. (2010) Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics* 2010;11:401.
- Cowen R.K., Paris C.B., Srinivasan A. (2006) Scaling of Connectivity in Marine Populations. *Science* 311:522-527.
- DeWoody J.A., Avise J.C. (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J Fish Biol* 56:461-473
- DeWoody J.A., Schupp J., Kenefic L., Busch J., Murfitt L., Keim P. (2004) Universal method for producing ROX-labeled size standards suitable for automated genotyping. *Biotechniques* 37:348-352
- Drummond A.J., Rambaut A., Shapiro B., Pybus O.G. (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22(5):1185-1192.
- Drummond A.J., Rambaut A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7:214.
- Dupanloup I., Schneider S., Excoffier L. (2002) A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* 11: 2571-2581.
- El Mousadik A., Petit R.J. (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theor. Appl. Genet.* 92:832-839.
- Excoffier L., Smouse P., Quattro J. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial restriction data.

- Genetics 131:479-491.
- Excoffier L., Laval G., Schneider S. (2005) ARLEQUIN ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform Online* 1:47–50
- Falush D., Stephens M., Pritchard J.R. (2003) Inference of population structure using multi-locus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.
- Felsenstein J. (1984) Distance methods for inferring phylogenies: a justification. *Evolution* 38:16-24.
- Franks J.S., Hoffmayer E.R., Comyns B.H., Hendon J.R., Blake E.M., Gibson D.P. (2007) Investigations of fishes that utilize pelagic *Sargassum* and frontal zone habitats in Mississippi marine waters and the northcentral Gulf of Mexico. Final Report. Gulf Coast Research Laboratory, Ocean Springs, Mississippi and U.S. Fish and Wildlife Service, Atlanta, Georgia
- Fu Y.-X. (1997) Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. *Genetics* 147:915-925.
- Fu Y.-X., Li W.-H. (1993) Statistical tests of neutral mutations. *Genetics* 133:693-709.
- Gold J.R., Saillant E., Cummings N.J., Renshaw M.A. (2011) Genetic Divergence and Effective Size among Lane Snapper in U.S. Waters of the Western Atlantic Ocean. *N. Am. J. Fish. Manage.*, 31:209-223.
- Goudet, J. 1995. F-STAT (Version 1.2): A computer program to calculate F-Statistics. *J. Hered.* 86:485-486.
- Goodyear C.P., Thompson N.B. (1993). An evaluation of data on size and catch limits for gray triggerfish in the Gulf of Mexico. NOAA/NMFS/SEFSC/Miami Lab. Contrib.
- Gower J.F.R., Hu C., Borstad G.A., Kings S. (2006) Ocean color satellites show extensive lines of floating *Sargassum* in the Gulf of Mexico. *IEEE Transactions on Geoscience and Remote Sensing*, 44, pp. 3619–3625.
- Gower J.F.R., Kings S. (2011) Distribution of floating *Sargassum* in the Gulf of Mexico and the Atlantic Ocean mapped using MERIS. *Int. J. Remote Sens.* 32:1917-1929.
- Gulland, J.A. (1965) Fish stock assessment: a manual of basic methods. John Wiley & Sons, NY.
- Hauser L., Adcock G.J., Smith P.J., Ramirez J.H.B., Carvalho G.R. (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proc. Natl. Acad. Sci. USA.* 99:11742–11747.
- Hilborn R., Walters C.J. (2001) Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, NY, NY, 570p.
- Hilborn R., Quinn T.P., Schindler D.E., Rogers D.E. (2003) Biocomplexity and fisheries sustainability. *Proc. Natl. Acad. Sci. Am.* 100:6564-6568.
- Ingram G.W. (2001) Stock structure of gray triggerfish, *Balistes caprisucus*, on multiple spatial scales in the Gulf of Mexico. Ph.D Dissertation, Department of Marine Sciences, University of South Alabama.
- Karlsson S., Saillant E., Gold J.R. (2009) Population structure and genetic variation of lane

- snapper (*Lutjanus synagris*) in the northern Gulf of Mexico. *Mar. Biol.* 156:1841-1855.
- Leblois R., Rousset F., Estoup A. (2004) Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics* 166: 1081–1092.
- Leblois R., Estoup A., Rousset F. (2009) IBDSim: a computer program to simulate genotypic data under isolation by distance. *Mol. Ecol. Resour.* 9: 107–109.
- Leis J.M. (1991) The Pelagic Stage of Reef Fishes: The Larval Biology of Coral Reef Fishes. Pages 183-230 in P.F. Sale, editor. *The Ecology of Fishes on Coral Reefs*. Academic Press.
- Lindenman K.C., Pugliese R., Waugh G.T., Ault J.S. (2000) Developmental pattern within a multispecies reef fishery: management implications for essential fish habitats and protected areas. *Bull. Mar. Sci.* 66:929-956.
- Mariano A.J., Ryan E.H., Perkins B.D., Smithers S. (1995) The Mariano Global Surface Velocity Analysis 1.0. USCG Report CG-D-34-95, 55 pp.
- Nei M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY, USA.
- Nunney L., Elam D.R. (1994) Estimating the effective population size of conserved populations. *Conserv. Biol.* 8:175-184.
- Peakall R., Ruibal M., Lindenmayer D.B. (2003) Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, *Rattus fuscipes*. *Evolution* 57:1182-1195.
- Peakall R., Smouse P.E. (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* 6:288-295.
- Pinsky M.L., Montes, Jr. H., Palumbi S.R. (2010) Using isolation-by-distance and effective density to estimate dispersal scales in anemonefish. *Evolution* 64:2688-2700.
- Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- Pritchard J.K., Wen X., Falush D. (2010). Documentation for structure software: Version 2.3. Available at http://pritchardlab.stanford.edu/structure_software/release_versions/v2.3.4/structure_doc.pdf (accessed 03/30/2014).
- Pruett C.L., Saillant E., Gold J.R. (2005) Historical population demography of red snapper (*Lutjanus campechanus*) from the northern Gulf of Mexico based on analysis of sequences of mitochondrial DNA. *Mar. Biol.* 147:593-602.
- Prugnolle F., De Meeus T. (2002) Inferring Sex-biased dispersal from population genetic tools: a review. *Heredity* 88: 161-165.
- Puebla O., Bermingham E., McMillan O. (2012) On the spatial scale of dispersal in coral reef fishes. *Mol. Ecol.* 21: 5675-5688.
- Raymond M., Rousset F. (1995) Genepop (version 1.2): a population genetics software for exact test and Ecumenicism. *J Hered* 86:248-249

- Reznick D.N., Ghalambor C.K. (2005) Can commercial fishing cause evolution: Answers from guppies (*Poecilia reticulata*). *Can. J. Fish. Aquat. Sci.* 62: 791-801.
- Rice W.R. (1989) Analyzing tables of statistical tests. *Evolution* 43: 223-225.
- Richards W.J., Linderman K.C. (1987) Recruitment dynamics of reef fishes: planktonic processes, settlement and demersal ecologies, and fisheries analysis. *Bull. Mar. Sci.* 41:392-410.
- Roberts C.M. (1997) Connectivity and management of Caribbean coral reefs. *Science* 278:1454-1457.
- Rousset F. (1997) Genetic Differentiation and Estimation of Gene Flow from FStatistics Under Isolation by Distance. *Genetics* 145:1219-1228.
- Rousset F. (2000) Genetic differentiation between individuals. *J. Evol. Biol.* 13:58–62.
- Rousset F. (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* 8, 103-106.
- Rousset F. (2008). Dispersal estimation: Demystifying Moran's I. *Heredity* 100: 231–232
- Rozas J., Sánchez-DelBarrio J.C., Messeguer X., Rozas R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and others methods. *Bioinformatics* 19:2496-2497.
- Rozen S., Skaletsky H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, pp 365-386
- Ruzzante D.E., Taggart C.T., Cook D. (1999). A review of the evidence for genetic structure of cod (*Gadus morhua*) populations in the NW Atlantic and population affinities of larval cod off Newfoundland and the Gulf of St Lawrence. *Fish. Res.* 43:79-97.
- Saillant E., Bradfield S.C., Gold J.R. (2010) Genetic variation and spatial autocorrelation among young-of-the-year red snapper (*Lutjanus campechanus*) in the northern Gulf of Mexico. *ICES J. Mar. Sci.* 67:1240-1250.
- Saillant E., Renshaw M.A., Cummings N.J., Gold J.R. (2012) Conservation genetics and management of yellowtail snapper (*Ocyurus chrysurus*) in the U.S. Caribbean and South Florida. *Fisheries Manage. Ecol.* 19:301-312.
- Sambrook J. E., Fritsch E. F., Maniatis T. (1989) *Molecular cloning: a laboratory manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- SEDAR-9 (Southeast Data, Assessment, and Review). 2006. Stock Assessment Report of SEDAR-9: Gulf of Mexico Gray Triggerfish. Assessment Report 1. Available: http://www.sefsc.noaa.gov/sedar/Sedar_Workshops.jsp?WorkshopNum=9.
- Simmons C.M., Szedlmayer S.T. (2011). Recruitment of Age-0 Gray Triggerfish to Benthic Structured Habitat in the Northern Gulf of Mexico. *Trans. Am. Fish. Soc.* 140:14-20
- Smouse P.E., Peakall R. (1999) Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. *Heredity* 82:561-573.
- Stephenson R.L. (1999) Sock complexity in fisheries management: a perspective of emerging issues related to population sub-units. *Fish. Res.* 43:247-249.

- Storz J., Beaumont M.A. (2002) Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution* 56:154-166
- Sunnucks P. (2000) Efficient genetic markers for population biology. *Trends Ecol. Evol.* 15: 199-203.
- Tamura K., Dudley J., Nei M., Kumar S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596-1599.
- Turner T.F., Wares J.P., Gold J.R. (2002) Genetic effective size is three orders of magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* 162:1329–1339.
- Ulanski S.L. (2008) *The Gulf Stream*. 2008. University of North Carolina Press, Chapel Hill.
- Valle M., Legault C.M., Ortiz M. (2001) A stock assessment for gray triggerfish, *Balistes caprisus*, in the Gulf of Mexico. Sustainable Fisheries Division Contribution SFD-00/01-124, NOAA Fisheries, Department of Commerce, Miami, FL.
- Van Oosterhout C., Hutchinson W.F., Wills D.P.M., Shipley P. (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535-538
- Waples R.S. (1995) Evolutionary significant units and the conservation of biological diversity under the Endangered Species Act. *American Fisheries Symposium* 17:8-27.
- Waples R.S. (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Hered.* 89:438–450.
- Waples R.S. (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. *Conserv. Genet.* 7:167-184.
- Waples R.S., Do C. (2010) Linkage disequilibrium estimates of contemporary N_e using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evol. Appl.* 3:244-262.
- Watts P.C., Rousset F., Saccheri I.J., Leblois R., Kemp S.J., Thompson D.J. (2007) Compatible genetic and ecological estimates of dispersal rates in insect (*Coenagrion mercuriale*: Odonata: Zygoptera) populations: analysis of ‘neighbourhood size’ using a more precise estimator. *Mol. Ecol.* 16: 737–751.
- Wells R.J.D., Rooker J.R. (2004) Spatial and temporal patterns of habitat use by fishes associated with *Sargassum* mats in the northwestern Gulf of Mexico. *Bull. Mar. Sci.* 74:81-99.
- Weir B.S., Cockerham C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-1370.
- Wilson C.A., Nieland D.L., Stanley A.L. (1995) Age, growth, and reproductive biology of gray triggerfish (*Balistes caprisus*) from the northern Gulf of Mexico commercial harvest. Final Report, Coastal Fisheries Institute, Louisiana State University.

Appendix Tables

Appendix Table 1. Characteristics of 21 microsatellite markers developed for gray triggerfish *Balistes capriscus*. Summary statistics are based on 35 specimens caught offshore the Louisiana coast (United States).

Locus	Repeat motif	Primer sequence 5'-3' (label)	T_a (°C)	A	Allele-size range	H_o/H_e	P_{HW}	GenBank accession no.
BC1	(TAGA) ₁₄	F:TGCCTGGAGGAGAAAAGAGA R:AATTCAAAACCAGGCCACAC (NED)	62	23	138-266	0.800/0.877	0.116	JQ285974
BC2	(GT) ₂₆	F:TACATGTGGGCAATGTGGTT (FAM) R:CTGGAATCTCCAGCATCTCA	62	22	150-200	0.886/0.948	0.308	JQ285975
BC3	(CA) ₃ CGTGGC G(CA) ₁₂ CT(CA) ₃	F:CATTGGTCTCCTTGCAAAA (FAM) R:ACAGACCCTGCTTTGCTTGT	60	5	146-156	0.514/0.599	0.702	JQ285976
BC9	(CA) ₁₃	F:GACGGATGAAACGTGAAGGT (FAM) R:CCTGGACCTGCTAAGTCTGC	60	22	204-256	0.971/0.951	0.149	JQ285977
BC13	(CA) ₁₅	F:GCGCCATTTTATTCTTGGAA (HEX) R:TGGTATGTGGCGGTCAATAA	60	21	227-277	0.85/0.892	0.799	JQ285978
BC14	(GT) ₃₂ GC(GT) ₄	F:CAGAGGAATCCACCTGATGA (FAM) R:CAGTCTCCAGATGGACCACA	60	18	101-197	0.829/0.918	0.053	JQ285979
BC16	(GT) ₁₂	F:CAGCAATGTGGGGTAAAGT (FAM) R:GCTTATGCGCTCAAAAAGTCC	58	4	249-257	0.486/0.503	0.508	JQ285980
BC17	(CA) ₁₈	F:TGGCTGGAATCAATGAACAA (NED) R:TTCACTCTGCTCATCCACA	56	12	205-253	0.629/0.608	0.513	JQ285981
BC19	(GT) ₁₄ CT(GT) ₂	F:TTGTCTCACCTCCCCTCTG (FAM) R:CTCTTGACATGCCACAAA	62	10	201-225	0.714/0.715	0.640	JQ285982
BC20	(CA) ₁₉	F:TCTGTGTTTCGGAACGTTT (FAM) R:TCCCGCCTTTTGTGTAGTTT	62	13	165-207	0.441/0.901	0.000	JQ285983
BC25	(GT) ₂₂	F:TCCCGCCTTTTGTGTAGTTT (HEX) R:TTCAGAAATGCTGCTGGATG	60	13	143-175	0.800/0.879	0.186	JQ285984
BC26	(CA) ₂₁	F:GGCGTAAAATCATCTAGGACA (HEX) R:TCAGCAAGCGTTCTCTACTA	62	16	188-244	0.886/0.877	0.985	JQ285985
BC27	(GT) ₁₃	F:TAAGTAACCGCAGCTCACC (HEX) R:ATACTCCACAGCCGAGACG	62	6	165-177	0.714/0.735	0.650	JQ285986
BC34	(GT) ₁₉	F:GCAGGTCACACTCCAGTC (HEX) R:TGTGCTCCCATCACACAGTT	62	14	186-272	0.829/0.874	0.336	JQ285987
BC36	(GT) ₁₂	F:GGGAACAAGGAGAGAGGAG (HEX) R:CTGGTGCCTGTGTGGTCTTA	62	7	264-278	0.629/0.733	0.003	JQ285988
BC41	(GT) ₁₂	F:TGGAATTGCTCGAGGCTAAG R:TGATATCAGCGACCTGTCA (HEX)	62	7	128-148	0.543/0.511	0.241	JQ285989
BC44	(CA) ₄ CG(CA) ₁₃	F:TTTGTCTGCAGAACCTCCT R:ACATTTGCAATCAGCAGCAC (FAM)	58	18	249-299	0.886/0.927	0.233	JQ285990
BC45	(CA) ₁₆ CG(CA) ₈	F:GCCTCAAGACCAGGAAGTG (FAM) R:TGTTTTGATCAGAGGCGTTG	54	10	251-271	0.829/0.810	0.553	JQ285991
BC46	(GT) ₂₈	F:TGAGAGTAAAATGGGTCACG R:TGAGCACGCTGGTTTGTTTA (FAM)	58	27	112-184	0.914 /0.967	0.227	JQ285992
BC47	(CA) ₂₈ TA(CA) ₄	F:ATCAGCTGCCTCCGTTTG (HEX) R:ATTAACCACCAGCCAACCTGC	56	26	108-174	0.886/0.961	0.066	JQ285993
BC49	(GT) ₁₄	F:TCATCAACTGTCCGGCTCTCA (FAM) R:CAACAGTCAGCCGTCTAGCA	56	7	244-270	0.618/0.502	0.933	JQ285994

Specific annealing temperature (T_a); number of alleles (A), observed heterozygosity (H_o), expected heterozygosity (H_e), probability of departure from Hardy-Weinberg equilibrium (P_{HW})

Appendix Table 2. Composition of multiplex panels developed for assay of gray triggerfish microsatellites.

<i>Multiplex #</i>	<i>Marker ID</i>	<i>Primer quantity (pmol/reaction)</i>	<i>Annealing temperature</i>
37	BC44	0.49	Touchdown 7 cycles Ta=62°C 7 cycles Ta=60°C 21 cycles Ta=58°C
	BC46	0.49	
	BC27	0.38	
	BC34	0.44	
	BC19	0.41	
39	BC47	0.57	56°C
	BC13	0.40	
	BC17	0.40	
	BC3	0.52	
	BC45	0.32	
40	BC14	0.73	60°C
	BC36	0.73	
	BC25	0.73	
44	BC26	0.60	Touchdown 7 cycles Ta=62°C 7 cycles Ta=60°C 21 cycles Ta=58°C
	BC41	0.60	
	BC2	0.60	
	BC49	0.40	
Simplex	BC16	2.2	58°C

Appendix Table 3a. Results of test of the 17 microsatellites developed for gray triggerfish in queen triggerfish (n = 32 samples).

<i>locus</i>	<i>Amplification issues</i>	<i>Scoring artifacts</i>	<i>A</i>	<i>Allele size range</i>	<i>H_e</i>	<i>F_{IS}</i>	<i>P_{H-W}</i>
BC13	1bp shifts	Null alleles	19	231-274	0.925	0.291	0.0026
BC14			3	89-105	0.552	0.155	0.1246
BC16			12	249-287	0.885	0.081	0.0776
BC17			7	212-224	0.757	0.092	0.4032
BC19			5	203-251	0.57	0.068	0.5296
BC2	Not scorable	-	-	-	-	-	-
BC25		Null alleles	9	139-165	0.881	0.539	0
BC26			17	221-277	0.868	0.064	0.0027
BC27			5	151-161	0.772	-0.093	0.2714
BC3			5	132-144	0.259	0.035	0.4805
BC34			17	233-273	0.909	0.038	0.3647
BC36			7	267-279	0.763	0.017	0.5129
BC41			10	136-154	0.87	-0.006	0.9534
BC44			8	239-261	0.497	-0.006	0.4687
BC45			7	242-256	0.591	-0.217	0.4168
BC46			5	103-125	0.584	-0.017	0.5705
BC47			22	112-164	0.914	-0.026	0.8556
BC49	No amplification	-	-	-	-	-	-

Number of alleles (*A*), expected heterozygosity (*H_e*), inbreeding coefficient (*F_{IS}*), probability of departure from Hardy-Weinberg equilibrium (*P_{H-W}*)

Appendix Table 3b. Results of test of the 17 microsatellites developed for gray triggerfish in ocean triggerfish (n = 16 samples).

<i>locus</i>	<i>Amplification issues</i>	<i>Scoring artifacts</i>	<i>A</i>	<i>Allele size range</i>	<i>H_e</i>	<i>F_{IS}</i>	<i>P_{H-W}</i>
BC13	No amplification	-	-	-	-	-	-
BC14	No amplification	-	-	-	-	-	-
BC16			3	251-255	0.327	0.236	0.4344
BC17			13	228-262	0.923	-0.016	0.762
BC19	No amplification	-	-	-	-	-	-
BC2	No amplification	-	-	-	-	-	-
BC25		Null alleles	7	153-169	0.717	0.39	0.0379
BC26	No amplification	-	-	-	-	-	-
BC27	No amplification	-	-	-	-	-	-
BC3	Monomorphic	-	-	148	-	-	-
BC34	No amplification	-	-	-	-	-	-
BC36			3	262-268	0.25	0.467	0.2031
BC41			5	160-172	0.488	-0.229	1
BC44	No amplification	-	-	-	-	-	-
BC45	Monomorphic	-	-	246	-	-	-
BC46	Monomorphic	-	-	96	-	-	-
BC47	Monomorphic	-	-	101	-	-	-
BC49	No amplification	-	-	-	-	-	-

Number of alleles (*A*), expected heterozygosity (*H_e*), inbreeding coefficient (*F_{IS}*), probability of departure from Hardy-Weinberg equilibrium (*P_{H-W}*)

Appendix Table 4. Haplotype distributions and summary statistics for gray triggerfish samples collected offshore from five localities along the U. S. coasts of the Gulf of Mexico and South Atlantic regions. TX: Texas, LA: Louisiana, WF: West Florida, EF: East Florida, SC: South Carolina.

Haplotype ID	Genebank accession number	TX	LA	WF	EF	SC
BC001	HQ599468	2	5	5	4	4
BC002	HQ599469	6	7	5	5	5
BC003	HQ599470	2	1	0	0	1
BC004	HQ599471	8	11	6	8	8
BC005	HQ599472	1	1	0	1	0
BC006	HQ599473	0	1	0	0	0
BC007	HQ599474	2	1	0	0	1
BC008	HQ599475	1	1	1	1	2
BC009	HQ599476	0	1	0	0	0
BC010	HQ599477	0	1	0	0	0
BC011	HQ599478	0	1	0	0	0
BC012	HQ599479	0	0	1	0	0
BC013	HQ599480	0	0	1	0	0
BC014	HQ599481	0	0	1	0	0
BC015	HQ599482	0	0	1	0	0
BC016	HQ599483	0	0	1	0	0
BC017	HQ599484	0	0	1	0	0
BC018	HQ599485	0	0	1	0	1
BC019	HQ599486	0	0	1	0	0
BC020	HQ599487	1	0	1	1	3
BC021	HQ599488	0	0	1	0	0
BC022	HQ599489	0	0	0	2	1
BC023	HQ599490	0	0	0	0	1
BC024	HQ599491	0	0	0	0	2
BC025	HQ599492	0	0	0	0	1
BC026	HQ599493	1	0	0	0	0
BC027	HQ599494	1	0	0	0	0
BC028	HQ599495	1	0	0	0	0
BC029	HQ599496	2	0	0	0	0
BC030	HQ599497	1	0	0	1	0
BC031	HQ599498	1	0	0	1	0
BC032	HQ599499	0	0	0	1	0
BC033	HQ599500	0	0	0	1	0
BC034	HQ599501	0	0	0	1	0
BC035	HQ599502	0	0	0	2	0
BC036	HQ599503	0	0	0	1	0
BC037	HQ599504	0	0	0	1	0
BC038	HQ599505	0	0	0	1	0
Number of haplotypes		16	14	11	12	15
Haplotype diversity		0.89195	0.81505	0.87957	0.90726	0.88736
Nucleotide diversity		0.00451	0.00345	0.00368	0.00403	0.00393

Appendix Table 5. Summary statistics for 17 nuclear-encoded microsatellites for gray triggerfish, *Balistes caprisacus*, sampled from various localities in the northern Gulf of Mexico. Localities (see also Figure 1): TX: South Texas, ETX: Galveston Bay, LA: Louisiana, MS: Mississippi, WF: West Florida, SWF: Southwest Florida, EF: East Florida, SC: South Carolina, FR: France, BE: Benin; N : sample size, A : number of alleles, A_R : allelic richness, H_E : gene diversity (expected heterozygosity), P_{HW} : probability of conforming to expected Hardy-Weinberg genotypic proportions, and F_{IS} : inbreeding coefficient measured as Weir and Cockerham's (1984) f .

Locus	TX	ETX	LA	MS	WF	SWF	EF	SC	FR	BE
<i>BC13</i>										
n	72	78	73	60	77	77	80	77	64	71
A	25	27	27	23	30	26	27	32	28	28
A_R	23.87	25.11	25.57	23.00	27.56	24.05	25.16	29.19	27.84	26.72
H_E	0.884	0.896	0.904	0.869	0.901	0.877	0.883	0.911	0.916	0.884
P_{HW}	0.155	0.094	0.173	0.920	0.908	0.215	0.958	0.025	0.626	0.191
F_{IS}	-0.021	0.027	0.015	-0.075	-0.037	0.008	-0.048	0.071	0.028	0.06
<i>BC14</i>										
n	72	78	74	61	77	77	80	76	64	71
A	24	25	24	24	26	21	27	26	25	26
A_R	23.39	23.88	23.17	23.90	24.50	20.00	25.66	24.34	24.87	25.21
H_E	0.924	0.930	0.910	0.928	0.903	0.917	0.930	0.919	0.910	0.853
P_{HW}	0.048	0.995	0.075	0.290	0.145	0.034	0.386	0.572	0.274	0.274
F_{IS}	0.023	0.021	0.109	0.046	0.051	0.065	0.006	-0.004	-0.03	0.025
<i>BC16</i>										
n	72	78	74	61	77	77	80	77	63	72
A	5	4	5	5	5	5	5	6	4	3
A_R	4.83	3.77	4.96	4.98	4.77	4.78	4.69	5.49	4.00	2.87
H_E	0.369	0.402	0.458	0.452	0.376	0.413	0.270	0.345	0.425	0.131
P_{HW}	0.781	0.282	0.496	0.593	0.4384	0.754	1.000	0.052	0.048	0.283
F_{IS}	-0.053	0.139	0.056	-0.015	-0.072	-0.037	-0.110	0.071	-0.194	0.153
<i>BC17</i>										
n	72	78	74	61	77	77	80	77	64	72
A	18	17	16	17	16	18	17	15	15	20
A_R	16.60	15.07	14.79	16.90	15.02	16.03	14.99	13.67	14.95	19.20
H_E	0.748	0.698	0.663	0.642	0.708	0.639	0.644	0.641	0.695	0.873
P_{HW}	0.534	0.212	0.045	0.226	0.063	0.491	0.012	0.264	0.677	0.038

F_{IS}	0.108	0.045	-0.060	-0.072	0.175	-0.017	0.126	0.020	0.011	0.030
----------	-------	-------	--------	--------	-------	--------	-------	-------	-------	-------

BC19

n	72	78	74	61	77	77	80	77	64	69
A	10	11	10	13	10	12	10	12	9	10
A_R	9.64	10.28	9.92	12.90	9.89	11.27	9.73	11.01	8.98	9.82
H_E	0.744	0.740	0.740	0.737	0.778	0.755	0.759	0.742	0.719	0.763
P_{HW}	0.080	0.343	0.753	0.773	0.077	0.169	0.370	0.268	0.851	0.880
F_{IS}	-0.138	-0.057	0.014	-0.023	0.015	0.106	-0.005	0.015	-0.066	-0.045

BC2

n	72	78	74	61	77	77	80	77	64	70
A	25	25	26	19	22	23	25	25	25	28
A_R	24.04	23.59	24.75	18.97	21.00	21.73	23.53	23.46	24.91	27.36
H_E	0.920	0.922	0.945	0.909	0.921	0.923	0.919	0.914	0.937	0.944
P_{HW}	0.315	0.288	0.003	0.458	0.403	0.405	0.347	0.143	0.650	0.764
F_{IS}	0.004	-0.001	0.056	0.026	-0.015	0.029	-0.007	0.018	0.032	-0.044

BC25

n	72	78	74	61	77	77	80	77	64	72
A	18	12	17	14	17	15	18	13	13	14
A_R	17.09	11.71	16.13	13.98	16.23	14.72	16.75	12.54	13.00	13.60
H_E	0.896	0.891	0.869	0.885	0.906	0.857	0.885	0.886	0.886	0.840
P_{HW}	0.575	0.774	0.241	0.356	0.580	0.496	0.639	0.811	0.308	0.700
F_{IS}	0.023	0.022	-0.010	-0.019	-0.047	0.0002	-0.073	-0.027	0.047	0.057

BC26

n	72	78	74	61	77	77	80	77	64	71
A	21	20	20	18	22	16	20	21	19	14
A_R	19.86	18.17	19.09	17.92	19.86	15.31	18.37	19.39	18.94	13.32
H_E	0.845	0.856	0.883	0.846	0.879	0.861	0.899	0.837	0.873	0.808
P_{HW}	0.769	0.280	0.960	0.108	0.665	0.158	0.744	0.890	0.420	0.183
F_{IS}	-0.035	0.072	-0.040	0.031	0.025	0.065	-0.015	0.004	0.087	-0.046

BC27

n	72	78	74	61	77	77	80	77	64	70
A	7	7	8	6	6	7	6	8	7	8
A_R	6.83	6.95	7.61	6.00	5.95	6.77	6.00	7.48	6.984	7.790
H_E	0.752	0.755	0.729	0.733	0.721	0.735	0.714	0.726	0.737	0.681

P_{HW}	0.733	0.489	0.588	0.111	0.566	0.710	0.468	0.279	0.779	0.985
F_{IS}	0.021	-0.018	0.017	0.128	-0.044	0.028	-0.033	-0.059	-0.017	-0.091

BC3

n	72	77	74	61	77	77	80	76	64	72
A	5	5	7	7	7	8	7	7	7	8
A_R	4.83	4.78	6.43	6.95	6.51	7.29	6.25	6.51	6.95	7.70
H_E	0.654	0.595	0.639	0.629	0.636	0.606	0.655	0.604	0.642	0.589
P_{HW}	0.066	0.550	0.600	0.103	0.209	0.516	0.146	0.170	0.487	0.935
F_{IS}	0.044	-0.047	0.1112	-0.042	-0.102	0.186	-0.011	-0.032	-0.095	-0.013

BC34

n	72	78	74	61	77	77	80	77	64	69
A	19	19	18	17	16	17	21	19	16	20
A_R	18.20	16.89	16.78	16.92	14.67	15.44	19.93	16.87	15.94	19.29
H_E	0.893	0.879	0.885	0.882	0.877	0.871	0.907	0.887	0.893	0.822
P_{HW}	0.274	0.014	0.302	0.105	0.544	0.674	0.269	0.192	0.468	0.540
F_{IS}	0.020	0.082	0.053	0.089	-0.036	0.016	-0.006	0.017	0.003	-0.004

BC41

n	72	78	74	61	77	77	80	77	64	72
A	8	7	8	9	9	9	9	9	9	4
A_R	7.81	6.94	7.77	8.95	8.71	8.63	8.48	8.48	8.98	3.87
H_E	0.571	0.54	0.512	0.532	0.547	0.569	0.534	0.569	0.631	0.540
P_{HW}	0.961	0.926	0.118	0.418	0.892	0.266	0.439	0.186	0.481	0.788
F_{IS}	-0.070	-0.068	0.023	0.015	0.027	0.018	0.016	-0.014	-0.015	0.049

BC44

n	72	78	74	61	77	77	80	77	64	72
A	22	26	23	24	27	21	24	23	23	23
A_R	21.58	24.08	21.96	23.90	25.06	20.22	22.58	21.71	22.91	22.06
H_E	0.937	0.938	0.926	0.941	0.939	0.931	0.935	0.920	0.937	0.920
P_{HW}	0.698	0.137	0.493	0.036	0.295	0.984	0.093	0.726	0.614	0.027
F_{IS}	0.007	0.016	0.066	0.006	0.046	-0.018	-0.030	-0.017	-0.034	0.049

BC45

n	72	78	74	61	77	76	80	77	64	72
A	12	13	11	13	13	12	14	13	13	12
A_R	11.33	12.47	10.96	12.97	12.32	11.56	13.15	12.51	12.97	11.57

H_E	0.795	0.813	0.821	0.843	0.828	0.827	0.833	0.807	0.826	0.774
P_{HW}	0.965	0.342	0.042	0.711	0.110	0.388	0.747	0.488	0.350	0.088
F_{IS}	-0.013	-0.089	-0.021	0.067	0.011	0.030	-0.006	-0.016	0.073	0.103

BC46

n	72	78	74	60	77	77	80	77	64	72
A	38	37	36	38	38	41	33	38	33	32
A_R	36.39	35.64	34.22	38.00	35.94	38.07	30.93	35.76	32.86	31.13
H_E	0.965	0.968	0.964	0.967	0.961	0.969	0.959	0.960	0.959	0.953
P_{HW}	0.746	0.407	0.092	0.928	0.071	0.169	0.262	0.059	0.373	0.926
F_{IS}	-0.008	-0.006	0.004	-0.017	0.040	0.008	0.036	0.039	0.023	-0.006

BC47

n	71	78	74	61	76	77	80	77	64	72
A	37	32	34	31	36	34	31	36	35	29
A_R	35.34	30.69	32.26	30.88	34.90	32.03	29.69	33.54	34.87	28.42
H_E	0.967	0.961	0.959	0.958	0.966	0.953	0.957	0.960	0.966	0.951
P_{HW}	0.629	0.495	0.477	0.907	0.424	0.867	0.598	0.561	0.115	0.339
F_{IS}	-0.005	-0.001	0.028	-0.044	0.020	-0.035	0.007	0.011	-0.002	-0.037

BC49

n	72	78	74	61	77	77	80	76	64	70
A	8	10	12	8	10	10	9	11	9	10
A_R	7.81	9.24	10.98	7.98	9.41	9.23	8.56	10.19	8.95	9.50
H_E	0.593	0.509	0.582	0.514	0.491	0.527	0.529	0.571	0.455	0.660
P_{HW}	0.876	0.735	0.240	0.217	0.754	0.903	0.634	0.254	0.191	0.003
F_{IS}	-0.101	-0.133	0.001	-0.052	-0.005	-0.010	-0.017	-0.0004	0.073	0.112

Appendix Table 6. Percentile distribution of the simulated distributions compatible with the isolation-by-distance slope estimated during the study.

<i>Distribution parameter</i>	<i>Percentile</i>			
	<i>25</i>	<i>50</i>	<i>75</i>	<i>90</i>
Pareto ($M=0.99$; $n=1.02$)	5	56	701	3396
Pareto ($M=0.96$; $n=0.602$)	261	1662	4769	7829
Geometric ($m=0.95$; $g=0.996$)	59	161	334	562
Geometric ($m=0.97$; $g=1$)	2268	4845	7422	8969
Sichel ($\gamma=-0.001$; $\xi=10000$; $\Omega=0.002$)	8	47	977	929
Sichel ($\gamma=-0.001$; $\xi=10000$; $\Omega=0.001$)	10	50	261	833

Appendix Table 7. Haplotype distributions and summary statistics for gray triggerfish samples collected offshore from five localities along the coasts of the United States (US), France (FR) and Benin (BE).

Haplotype ID	Genbank accession number	Haplotype count			Haplotype frequency (%)			
		US	FR	BE	US	FR	BE	Average
BC001	HQ599468	20	5	2	13.33	13.51	5.13	10.66
BC002	HQ599469	28	6	0	18.67	16.22	0.00	11.63
BC003	HQ599470	4	1	0	2.67	2.70	0.00	1.79
BC004	HQ599471	41	14	2	27.33	37.84	5.13	23.43
BC005	HQ599472	3	1	0	2.00	2.70	0.00	1.57
BC006	HQ599473	1	0	0	0.67	0.00	0.00	0.22
BC007	HQ599474	4	0	0	2.67	0.00	0.00	0.89
BC008	HQ599475	6	0	0	4.00	0.00	0.00	1.33
BC009	HQ599476	1	0	0	0.67	0.00	0.00	0.22
BC010	HQ599477	1	1	0	0.67	2.70	0.00	1.12
BC011	HQ599478	1	0	0	0.67	0.00	0.00	0.22
BC012	HQ599479	1	0	0	0.67	0.00	0.00	0.22
BC013	HQ599480	1	0	0	0.67	0.00	0.00	0.22
BC014	HQ599481	1	0	0	0.67	0.00	0.00	0.22
BC015	HQ599482	1	0	0	0.67	0.00	0.00	0.22
BC016	HQ599483	1	0	0	0.67	0.00	0.00	0.22
BC017	HQ599484	1	0	0	0.67	0.00	0.00	0.22
BC018	HQ599485	2	0	0	1.33	0.00	0.00	0.44
BC019	HQ599486	1	1	0	0.67	2.70	0.00	1.12
BC020	HQ599487	6	1	29	4.00	2.70	74.36	27.02
BC021	HQ599488	1	0	0	0.67	0.00	0.00	0.22
BC022	HQ599489	3	0	0	2.00	0.00	0.00	0.67
BC023	HQ599490	1	0	0	0.67	0.00	0.00	0.22
BC024	HQ599491	2	0	0	1.33	0.00	0.00	0.44
BC025	HQ599492	1	0	0	0.67	0.00	0.00	0.22
BC026	HQ599493	1	0	0	0.67	0.00	0.00	0.22
BC027	HQ599494	1	0	0	0.67	0.00	0.00	0.22
BC028	HQ599495	1	0	0	0.67	0.00	0.00	0.22
BC029	HQ599496	2	0	0	1.33	0.00	0.00	0.44
BC030	HQ599497	2	0	0	1.33	0.00	0.00	0.44
BC031	HQ599498	2	0	0	1.33	0.00	0.00	0.44
BC032	HQ599499	1	0	0	0.67	0.00	0.00	0.22
BC033	HQ599500	1	1	0	0.67	2.70	0.00	1.12
BC034	HQ599501	1	0	0	0.67	0.00	0.00	0.22
BC035	HQ599502	2	1	0	1.33	2.70	0.00	1.35
BC036	HQ599503	1	0	0	0.67	0.00	0.00	0.22

BC037	HQ599504	1	0	0	0.67	0.00	0.00	0.22
BC038	HQ599505	1	0	0	0.67	0.00	0.00	0.22
BC039	KJ670378	0	1	0	0.00	2.70	0.00	0.90
BC040	KJ670379	0	1	0	0.00	2.70	0.00	0.90
BC041	KJ670380	0	1	0	0.00	2.70	0.00	0.90
BC042	KJ670381	0	1	0	0.00	2.70	0.00	0.90
BC043	KJ670382	0	1	0	0.00	2.70	0.00	0.90
BC044	KJ670383	0	0	1	0.00	0.00	2.56	0.85
BC045	KJ670384	0	0	1	0.00	0.00	2.56	0.85
BC046	KJ670385	0	0	1	0.00	0.00	2.56	0.85
BC047	KJ670386	0	0	1	0.00	0.00	2.56	0.85
BC048	KJ670387	0	0	1	0.00	0.00	2.56	0.85
BC049	KJ670388	0	0	1	0.00	0.00	2.56	0.85
Sample size					150	37	39	
Number of haplotypes					38	15	9	
Haplotype diversity					0.871	0.826	0.449	
Nucleotide diversity					0.0040	0.0035	0.001	