

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

DR. ERIC SAILLANT (Orcid ID : 0000-0002-1480-4515)

Article type : Original Article

Spatial connectivity in an adult-sedentary reef fish with extended pelagic larval phase

L. Antoni and E. Saillant

The University of Southern Mississippi
School of Ocean Science and Technology
Gulf Coast Research Laboratory, 703 East Beach Dr.
Ocean Springs, MS 39564

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/mec.14263](https://doi.org/10.1111/mec.14263)

This article is protected by copyright. All rights reserved

20 **Keywords:** *Balistes capriscus*, gray triggerfish, dispersal, connectivity, population genetics-
21 Empirical

22

23

24 **Corresponding author:**

25

26 Eric Saillant

27 School of Ocean Science and Technology

28 University of Southern Mississippi, Gulf Coast Research Laboratory

29 Ocean Springs, MS 39564 USA

30 Fax: 228-872-4204

31 E-mail: eric.saillant@usm.edu

32

33 **Running title:** Spatial connectivity in gray triggerfish

34

35 **Abstract**

36 Understanding the spatial scale of demographic connectivity in marine reef fishes dispersing
37 pelagic larvae is a challenging task because of the technical difficulties associated with tagging
38 and monitoring the movements of progeny at early life stages. Several studies highlighted a
39 strong importance of local retention with levels of dispersal of ecological significance restricted
40 to short distances. To date little information is available in species where pelagic dispersal lasts
41 for long periods of time. In this work population structure and connectivity were studied in the
42 gray triggerfish, *Balistes capriscus*. Gray triggerfish larvae and juveniles remain associated with
43 floating *Sargassum sp.* beds for an estimated period of 4 to 7 months before settling on benthic
44 habitats where they remain sedentary as adults. Analysis of genetic variation among populations
45 along the continental shelf of the northern Gulf of Mexico and U.S. east coast, encompassing
46 over 3,100 km of coastline, revealed homogeneous allele frequencies and a weak isolation by
47 distance pattern. Moment and maximum-likelihood estimates of dispersal parameters both
48 indicated occurrence of large neighborhoods with estimates of the dispersal distribution

49 parameter σ of 914 and 780 km respectively. Simulated distributions of dispersal distances using
50 several distribution functions all featured substantial fractions of long distance dispersal events
51 with the 90% percentiles of travel distance prior to settlement averaging 1,809 km. These results
52 suggest a high dependency of local recruitment on the output of non-local spawning stocks
53 located hundreds of kilometers away and a reduced role of local retention in this species.

54
55
56

57 **Introduction**

58 Characterizing genetic and demographic connectivity among geographic populations is
59 essential to design effective conservation strategies (Lowe & Allendorf 2010). The marine
60 environment is *a priori* open to migrations (Awise 1998) and many marine species display a
61 continuous distribution across large portions of their range, sometimes encompassing several
62 thousand kilometers, leading to the assumption that connectivity occurs across these large
63 geographic areas, promoted by free dispersal. The spatial scale of the actual genetic connectivity
64 is however influenced by several factors including the dispersal capability of organisms, the
65 density of populations and the strength of local adaptation. The dispersal itself is determined by
66 several factors including the occurrence of barriers to gene flow resulting from discontinuities of
67 suitable habitat, the duration of the physical transport of eggs and larvae, the velocity of currents
68 involved for species with pelagic planktonic phases, and the movement behavior and capabilities
69 of adults. In reef fishes, movements of adults are often limited, and when this is the case, the
70 larval transport processes are assumed to be the major determinants of dispersal (Leis &
71 McCormick 2002; Jones *et al.* 2009; Shanks 2009).

72 Under the island model, genetic connectivity can be maintained even when only a few
73 effective migrants are exchanged per generation (Waples 1998) which is often enough to rapidly
74 ensure the spread of advantageous mutations across a metapopulation (Lowe & Allendorf 2010).
75 However, management is also concerned with local demographic change of populations, in
76 particular the relative role of local recruitment and migration in determining local demographic
77 dynamics, or the potential for local replenishment through migration from external populations
78 (Kritzer & Sale 2004). The spatial scale of this demographic connectivity is often different from

79 that of the genetic connectivity and is also more difficult to determine because it requires
80 estimating rates of migrations. Obtaining direct estimates of the spatial scale of demographic
81 connectivity requires data on local recruitment as well as quantitative estimates of migrations
82 from and to other demes (Lowe & Allendorf 2010). This information is particularly challenging
83 to obtain when boundaries between demes are not clearly defined as is the case in many marine
84 species that are structured in large continuous metapopulations. In such cases, tag and recapture
85 studies or studies of elemental signatures in otoliths can provide information on juvenile and
86 adult movements but are not adapted to measure dispersal in most reef fishes that are sedentary as
87 adults but disperse planktonic eggs and larvae that cannot easily be tagged (Thorrold *et al.* 2002).
88 Particle tracking may be used to predict larval envelopes (e.g. Roberts 1997; Cowen *et al.* 2006;
89 Johnson *et al.* 2009) but this approach can also be challenging in species that cannot be modelled
90 by a simple particle including, for example, those utilizing pelagic habitats that are fluctuating
91 over time in size and shape such as floating *Sargassum* beds.

92 Paternity analysis inferred from molecular marker data and tracking of the maternal origin of
93 settling juveniles through the analysis of stable isotopes transmitted from mother to offspring
94 have been used successfully to demonstrate occurrence of local recruitment (e.g. Christie *et al.*
95 2010, Almany *et al.* 2007) but these approaches are limited when populations are large and
96 dispersal occurs across broad geographic areas. Genetic estimation of contemporaneous rates of
97 gene flow through assignment tests has been used in several species (Lowe & Allendorf 2010)
98 but this approach requires migrants to be exchanged between discrete and differentiated
99 populations. When there is isolation-by-distance in a continuous population this method is
100 irrelevant but inferences on dispersal can be made using the isolation-by-distance theoretical
101 framework (Rousset 1997; Puebla 2009). Recent developments of this approach using individual
102 models and maximum likelihood algorithms (Watts *et al.* 2007; Rousset & Leblois 2007, 2012)
103 allowed assessing dispersal in metapopulations showing high degree of genetic connectivity (and
104 homogeneity) across large geographic areas (e.g. Puebla *et al.* 2012).

105 Studies in reef fishes to date have revealed relatively small (less than 100 km in most cases)
106 larval dispersal envelopes (Roberts 1997; Cowen *et al.* 2006; Shanks *et al.* 2009; Puebla *et al.*
107 2012), but the species considered were characterized by short dispersal durations, usually less
108 than a month. On another hand, data on the spatial scale of demographic connectivity are lacking

109 for species where larval dispersal lasts longer. In those species, rare successful long distance
110 dispersal events could maintain genetic connectivity across long distances even if the majority of
111 dispersal events are restricted to local areas; in that situation, the local spawning biomass would
112 retain a strong influence on recruitment. Alternatively, longer larval transport could result in high
113 proportion of dispersal events at long distances and a reduced contribution of local spawning
114 stocks to recruitment. Distinguishing between these scenarios is essential in order to determine
115 effective conservation and management strategies.

116
117 The gray triggerfish, *Balistes caprisacus*, is a reef fish that inhabits sub-tropical and temperate
118 waters on both sides of the Atlantic Ocean. This species is highly sedentary as adult where it is
119 found associated to benthic structures of the continental shelf (Ingram 2001) at depths ranging
120 between 0 and 100 m (Harmelin-Vivien & Quérou 1990). Dispersal is thought to occur primarily
121 during the larval and juvenile stages (Wells & Rooker 2004; Franks *et al.* 2007) when the species
122 is pelagic. This pelagic phase (4-7 months, Simmons 2008) lasts longer than in most other reef
123 fishes, and, during that period, larvae and juveniles are found associated with floating seaweeds
124 and flotsam (mostly *Sargassum sp.*) until they settle on hard benthic structures. Gray triggerfish
125 reach sexual maturity at a length of 250 mm fork length (FL) and the age of 1 year for males and
126 2 years for females (Wilson *et al.* 1995; Ingram 2001). Females produce on average 13,809
127 oocytes per gram of ovary (range 6,318 - 24,188, Hood & Johnson 1997). Gray triggerfish can
128 live up to 16 years in the Gulf of Mexico (NMFS 2006) and their generation time is estimated
129 between 4 and 8 years (Jing *et al.* 2015). Their center of abundance is located in the southeast
130 United States (Gulf of Mexico and southeast U.S. coast) where they approach a continuous
131 distribution along shelf habitats. The life history features of this species predict structuring
132 according to an isolation-by-distance model as discussed above where dispersal is limited by the
133 spatial scale of the pelagic larval transport. The extended pelagic phase could promote long
134 distance movement but it has also been hypothesized that larvae could be retained in local eddies
135 and recruit close to their spawning location (NMFS 2006). The availability of a large
136 continuously distributed population in the southeastern U.S. provides the opportunity to describe
137 the isolation-by-distance model and assess quantitatively the spatial scale of demographic
138 connectivity resulting from larval dispersal in this species.

139 In this work, 17 microsatellite markers were used to survey genetic variation among gray
140 triggerfish in the northern Gulf of Mexico and along the east coast of the United States. The
141 dataset was used to characterize patterns of population structure and provide first estimates of
142 dispersal parameters and connectivity in the species.

143

144

145 **Materials and methods**

146

147 **Sampling**

148 Samples of gray triggerfish were obtained during the summer and fall of 2008, 2009 and
149 2010. Sampling focused on sub-adult and adult specimens settled on benthic continental shelf
150 habitats where they are known to display high site fidelity (Ingram 2001). Specimens from across
151 the northern Gulf of Mexico were obtained in conjunction with the summer and fall groundfish
152 SEAMAP surveys conducted by the National Marine Fisheries Service (NOAA-Fisheries). The
153 survey employs a stratified randomized design to sample benthic shelf habitats used by triggerfish
154 juveniles and adults, (10-100 m depth) by trawling from Pensacola to the U.S./Mexico border
155 (Nichols 2004). Additional samples from the northern Gulf were collected at recreational fishing
156 docks (Mississippi and vicinity of Panama City, Florida) and during fishery-independent reef fish
157 monitoring surveys conducted by the NOAA-Fisheries Panama City laboratory in west Florida
158 (east of Pensacola) using traps. Sampling in the northern Gulf (1,400 km of coastline, Figure 1)
159 yielded 430 specimens and resulted in minimal gaps in this section of the studied range except for
160 the shelf nearing the Mississippi estuary delta and a small portion of the Texas shelf north of
161 Corpus Christi.

162 Two hundred and thirty-five additional samples were obtained from southwest Florida (SWF,
163 $n = 77$ collected by trawling), from southeast Florida (SEF, $n = 80$ collected by angling) and
164 South Carolina (SC, $n = 78$ collected by trapping) bringing the total sampling size to 665
165 specimens.

166 Specimens were preserved frozen on board (SEAMAP samples) or kept on ice until fish were
167 landed. Muscle tissue and fin clips were collected and stored in 95% alcohol or a Dimethyl
168 Sulfoxide (DMSO) salt-saturated storage buffer (0.25 M EDTA, 20% DMSO, 30% H₂O, and

169 NaCl) prior to DNA extraction except for the samples from South Carolina which were preserved
170 in a Sarkosyl urea lysis buffer (1% *N*-lauroylsarcosinate, 20 mM NaPO₄, 8 M urea, 1 mM
171 EDTA).

172
173 Laboratory assays
174 DNA extraction was performed following a phenol-chloroform protocol (Sambrook *et al.*
175 1989). The fish were genotyped at 17 microsatellite markers described in Antoni & Saillant
176 (2012). To improve the cost effectiveness of genotyping, microsatellites were assayed in four
177 multiplex panels developed during the study. Detailed multiplexed PCR protocols including
178 microsatellite loci identification, primers concentration, fluorescent labeling, and specific T_a are
179 presented in Appendix 1. PCR products were loaded on a 6% acrylamide gel and run on an ABI
180 Prism 377 DNA Sequencer (Applied Biosystems, Foster City, CA, USA) following instructions
181 from the manufacturer. Electropherograms were analyzed in the software GENESCAN v.3.1.2
182 (Applied Biosystems) and alleles were called in the software GENOTYPER v.2.5 (Applied
183 Biosystems).

184
185 Data analysis
186 Samples were initially grouped in six regional populations based on gaps in sampling
187 (Figure 1). The occurrence of scoring errors due to null alleles, stuttering bands, and large allele
188 dropout in each regional population-sample was tested in MICROCHECKER v.2.2.3 (Van
189 Oosterhout *et al.* 2004). The conformance of genotype proportions to Hardy-Weinberg (H-W)
190 equilibrium expectations was tested using exact tests in GENEPOP v.4.2 (Raymond & Rousset
191 1995; Rousset 2008a). Probability-value estimates were based on 10,000 dememorizations, 500
192 batches, and 5,000 iterations per batch. Departure from H-W equilibrium (F_{IS}) measured as Weir
193 and Cockerham's (1984) f , the number of alleles, allelic richness (El Mousadik & Petit 1996), and
194 gene diversity (expected heterozygosity calculated as described in Nei 1987) were computed for
195 each regional sample in FSTAT v.2.9.3 (Goudet 1995).

196
197 Analysis of spatial genetic variation

198 Homogeneity in allelic richness and gene diversity among samples was tested using the Friedman
199 ranks test, as implemented in SPSS v.20 (IBM Corp., Armonk, NY, USA). The degree of
200 population differentiation (F_{ST}) among regions was estimated as Weir and Cockerham (1984) θ
201 as calculated in FSTAT and homogeneity of allele distributions among regional samples was tested
202 using exact tests in GENEPOP. Pairwise comparisons were performed by computing estimates of
203 pairwise θ between individual regions and performing associated pairwise exact homogeneity
204 tests. Markov Chain parameters during exact homogeneity tests were the same as above (Exact
205 tests of H-W equilibrium). The False Discovery Rate (FDR, Benjamini & Hochberg 1995)
206 procedure was used to determine the significance threshold for P-values when multiple
207 independent tests were conducted simultaneously.

208 Isolation-by-distance due to limited dispersal potential and barriers to gene flow (genetic
209 discontinuities) may both account for divergence among geographic samples. Spatial genetic
210 variation within the region was therefore further explored using the Bayesian clustering approach
211 implemented in the software TESS v.2.3.1 (Chen *et al.* 2007; Durand *et al.* 2009a). TESS aims to
212 detect genetic discontinuities within continuously distributed populations of a species based on
213 the distribution of multilocus genotypes. This approach accounts for the decay of spatial
214 autocorrelation that occurs due to isolation-by-distance, and is therefore well suited for
215 populations displaying spatially restricted dispersal and a predicted isolation-by-distance pattern.
216 One hundred runs were performed using a conditional autoregression (CAR) admixture model,
217 allowing for correlated allele frequencies among populations. Each Monte Carlo simulation
218 included 250,000 sweeps with the first 50,000 sweeps discarded as burn-in. The 20 runs showing
219 the lowest Deviance Information Criteria (Spiegelhalter *et al.* 2002) were retained to make
220 inferences, as recommended by Durand *et al.* (2009b).

221 Structuring according to an isolation-by-distance mechanism was examined within ranges
222 where no evidence of genetic discontinuity was found. The method developed by Rousset (2000)
223 and Leblois *et al.* (2004) was employed as it allows estimating dispersal parameters based on
224 existing theory of isolation-by-distance (Rousset 1997).

225 The genetic distance between pairs of individuals was estimated as the \hat{e} statistics (Watts *et*
226 *al.* 2007) computed in the software GENEPOP. The \hat{e} statistics is more powerful in cases where the

227 spatial pattern of population structure is weak (Watts *et al.* 2007), as is the case in the present
228 study (see results section). The analysis of isolation-by-distance focused on data obtained on
229 specimens ($n = 430$) collected between south Texas and west Florida (1,400 km) because this
230 portion of our sampling design approached best a continuous sampling along the coastline as
231 recommended to infer parameters of the model (Leblois *et al.* 2004). Considering the shelf habitat
232 used by gray triggerfish, two approaches were used to compute individual coordinates and
233 calculate geographic distances between individuals and isolation-by-distance statistics. In a first
234 approach a one dimensional lattice (mid-shelf transects following the coastline, 1D model) was
235 used thus assuming dispersal in a one dimensional linear habitat. In a second approach, a two
236 dimensional habitat spanning from Texas to west Florida was considered (2D model). The 2D
237 model could only be evaluated using the likelihood approach in MIGRAINE where the shape of the
238 2D lattice could be specified.

239 Because estimation of the parameters of the isolation-by-distance model is biased when the
240 geographic distance between samples being compared is greater than $0.56\sigma/\sqrt{2\mu}$, where σ is the
241 standard deviation of parental position relative to offspring position and μ is the mutation rate
242 (Rousset 1997), a bootstrap resampling approach was used to investigate the effect of the spatial
243 scale of sampling on estimates of the slope of the isolation-by-distance relationship and σ .
244 Subsamples were drawn by resampling sets of 100 individuals located within subsections of the
245 lattice of various lengths using the software POPTOOLS v.3.2.5 (Hood 2010) and the slope of the
246 1D linear regression between genetic and geographic distance (b) was estimated for each
247 resampled dataset.

248 This slope was then used to calculate σ , given the effective population density (D), using the
249 relationship (Rousset 1997)

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

251
252 Inferences on σ thus require information on population density. Two approaches were taken
253 to obtain values for D and discuss values of σ and the distribution of dispersal distances. An
254 upper bound for D is given by the census population density (D_c). The census density of gray

255 triggerfish was estimated based on average landing data in the Gulf of Mexico during the
256 sampling period obtained from the recreational fisheries statistics database of the Fisheries
257 Statistics Division of the National Marine Fisheries Service (personal communication, database
258 accessed 08 January 2016) and accounting for estimates of fishing mortality rates that range
259 between 0.435 and 0.53 (NMFS 2011). The obtained estimate of the census number of adults was
260 applied to the Gulf section of the lattice (2,035 km) to derive estimates of census density for the
261 1D model and to estimates of the area of the shelf habitat for gray triggerfish approximated as a
262 strip surrounding the 1D lattice for the 2D model.

263 Effective density (D_e) was also estimated using genetic data. Considering the observed
264 homogeneity in allele frequencies across the sampling surface (see results), an estimate of the
265 effective size for the overall metapopulation was generated using the maximum likelihood (ML)
266 coalescent approach in the software MIGRAINE v.0.4.1 (Rousset & Leblois 2007, 2012; Leblois *et*
267 *al.* 2014). The OnePopVarSize demographic model allowing accounting for historical change in
268 population size was used in the estimation (Appendix 2). The parameter N that represents an
269 estimate of the current effective population size was calculated assuming an average mutation
270 rate across microsatellites of 5×10^{-4} (Estoup & Angers 1998). N was also calculated considering
271 mutation rates of 10^{-3} and 10^{-4} in order to evaluate the sensitivity of parameter estimates to the
272 mutation rate. The obtained estimate of N was applied to the entire lattice length/surface to derive
273 an estimate of effective density.

274 Estimates of contemporaneous N_e by the linkage disequilibrium method were also generated
275 for each of the 6 regional populations using the software LDNE (Waples & Do 2010).

276
277 Because the genetic consequences of dispersal depend on the shape of the distribution of
278 dispersal distance (Rousset 2008b), a simulation approach after Puebla *et al.* (2012) was taken to
279 determine the parameters of dispersal distance distributions yielding isolation-by-distance slopes
280 consistent with that estimated from the empirical dataset. Coalescent simulations were
281 implemented in the software IBDSIM v.2.0 (Leblois *et al.* 2009) considering various distribution
282 functions (Geometric, Pareto, and Sichel). Simulations employed a one dimensional lattice of
283 10,000 km with absorbing boundaries; samples were generated from a 1,400 node subsection of
284 the lattice to match the length of the portion of the northern Gulf of Mexico (south Texas to west

285 Florida, approximately 1,400 km of coastline) used in the empirical study and at most one
286 individual was sampled per node. Simulated datasets included 17 unlinked loci following a GSM
287 mutation model with a mean mutation rate of 5×10^{-4} and a geometric variance of multi-step
288 mutations with parameter estimated during N_e estimation in MIGRAINE (Appendix 2). The
289 simulated datasets were processed for isolation-by-distance analysis as described above.
290 Parameters for each of the dispersal distribution functions were adjusted to determine ranges of
291 values leading to isolation-by-distance slopes b similar to those obtained with the empirical
292 dataset. Series of simulations were then conducted in triplicates within this range to identify the
293 parameter values (or combination of parameter values) that led to isolation-by-distance slopes
294 closest to the estimates from the empirical dataset. The influence of the mutation rate on the
295 dispersal distribution parameters was evaluated by considering mutation rates of 10^{-3} and 10^{-4}
296 used as an upper and lower bound of the average mutation rate for the 17 microsatellites used in
297 the study, respectively.

298 Finally, Maximum Likelihood estimates of σ were generated using both linearIBD and
299 planarIBD demographic models implemented in MIGRAINE. These methods provide an estimate
300 of the neighborhood size parameter (Nb) from which an estimate of σ can be derived. The
301 planarIBD model accounts for a two dimensional habitat while the linearIBD model assumes
302 dispersal along a one dimension (linear) lattice. Estimates were generated during three replicate
303 runs employing the Product of Approximate Conditional (PAC) likelihoods algorithm with 2,000
304 points and 100 runs per point.

305 Estimates of the parameter σ were derived from Nb using the relationships $Nb = 2D\sigma^2$
306 (equation 2) and $Nb = 2D\pi\sigma^2$ (equation 3) for the linear and the two dimensional model
307 respectively where D was set to the census or the effective population density value determined
308 as above.

309 An exclusion approach in the software GENECLASS v.2.0 (Piry *et al.* 2004) was used to test
310 the influence of possible migrants from divergent gray triggerfish populations on estimates of
311 isolation-by-distance parameters. Sampled individuals were assigned to a locality based on their
312 multilocus genotype using the Bayesian method of Rannala & Mountain (1997); the probability
313 that an individual belonged to a given locality was calculated using the resampling algorithm of
314 Paetkau *et al.* (2004) and was based on 10,000 simulated individuals. Putative migrants were

315 identified as those showing P-value of assignment below 0.05 for all 6 regional samples. The
316 slope of the isolation-by-distance model and sigma were re-calculated as described above after
317 removing the detected possible migrants from the dataset.

318

319 **Results**

320 Four out of 102 tests (6 geographic samples x 17 loci) of Hardy-Weinberg equilibrium were
321 significant before FDR correction for multiple tests performed simultaneously. None of the test
322 remained significant after correction. MICROCHECKER analyses indicated possible occurrence of
323 null alleles at locus BC14 in the ETX-LA region, locus BC17 in the SEF region, and stuttering
324 and/or null alleles at locus BC3 in the SWF region. Because the scoring artifacts at these three
325 loci were found in one region (out of 6) only and did not lead to significant departure from
326 Hardy-Weinberg expectation, all 17 markers were kept for further analysis.

327 Summary statistics per locus and per region including number of alleles, allelic richness, gene
328 diversity, inbreeding coefficient, and probability of significance of tests of Hardy-Weinberg
329 equilibrium are presented in Appendix 3. The number of alleles (A) per locus averaged 25.6 and
330 ranged between 9 (locus BC16) and 45 (locus BC46). Gene diversity ranged between 0.27 (locus
331 BC16 in the SEF region) and 0.969 (locus BC46 in the SWF region). Allelic richness and gene
332 diversity did not differ significantly among localities ($P = 0.240$ and $P = 0.083$ respectively).

333 The estimate of θ was very low (0.0004, 95% bootstrapping Confidence Interval CI: 0-0.001)
334 and the probably that θ differed from zero from exact homogeneity tests was 0.031. Homogeneity
335 tests at individual loci did not reveal significant heterogeneity in allele frequencies among regions
336 except for one locus, BC46, that showed significant heterogeneity ($P = 0.042$) before FDR
337 correction but not after correction. Pairwise θ values between individual regions averaged 0.0006
338 (range -0.0006-0.0018, Table 1) and only two pairwise exact homogeneity tests (across loci) were
339 significant before and after FDR correction (SWF versus ETX-LA comparison: $P = 0.0177$,
340 estimate of $\theta = 0.0008$; SWF versus SEF comparison: $P = 0.0032$, estimate of $\theta = 0.0018$).
341 Bayesian clustering runs in TESS all converged towards a single unit with no genetic
342 discontinuity within the sampled range. Further analysis of isolation-by-distance proceeded under
343 this assumption.

344 The estimate of the current effective size (N) derived assuming an average mutation rate
345 of 5×10^{-4} was 29,940 individuals (95% CI: 18,570-62,630, Appendix 2). The genetic estimate of
346 D_e was generated by applying the estimate of N to the entire one dimensional lattice (from south
347 Texas to South Carolina, 3,100 km) yielding a value of 9.66 individuals/km (95% CI: 5.99-20.20)
348 for D_e for the 1D model. N was then applied to a 20-km wide strip surrounding the transect line
349 ($123,331 \text{ km}^2$) for the 2D model yielding a value of 0.24 ind./ km^2 (95% CI: 0.15-0.51) providing
350 an upper bound value for D_e under this model. The census density D_c for the Gulf of Mexico was
351 175 ind./km (1D model) or 2.89 ind./ km^2 (2D model) giving a ratio of effective to census density
352 of 0.055 (95% CI: 0.034-0.115) for the 1D model and 0.083 (95% CI: 0.052-0.176) for the 2D
353 model.

354 All the obtained N_e estimates from the linkage disequilibrium method were infinite or very
355 large (greater than 2,494, Appendix 4).

356 Estimates of the isolation-by-distance slope under the 1D model using subsets of the data
357 encompassing increasing distance ranges revealed a high variance among slopes when resampled
358 datasets were generated using genotypes found within short distance ranges ($< 1,100 \text{ km}$,
359 Appendix 5). The mean and standard error of slopes from resampled datasets stabilized between
360 3.4×10^{-8} and 4.4×10^{-7} when the sampled range was between 1,400 and 1,700 km (Appendix 5).
361 Accordingly, final estimates were generated based on all available data for the area between south
362 Texas and west Florida where the high density of sampling locations with minimal gaps best
363 reflected the near-continuous distribution of gray triggerfish along the continental shelf. The
364 obtained estimate was 3.1×10^{-8} (lower and upper bounds of the slope -5.24×10^{-7} and 4.61×10^{-7} ,
365 Figure 2). Point estimates generated using greater portions of the dataset (i.e. including
366 localities in south Florida and South Carolina) were all included within the bounds of the
367 confidence interval described above. Considering the genetic estimate of effective density and
368 census density, the corresponding values of σ derived using equation 1 were 914 (95% CI: 237-
369 $+\infty$) and 215 (95% CI: 56- $+\infty$) respectively.

370 Estimates of σ derived from Nb values obtained from the maximum likelihood approach in
371 MIGRAINE using equations 2 and 3 were 780 (95% CI: 255-2517) for the 1D model and 740 (95%
372 CI: NA-7,330) for the 2D model. Because dispersal along a coastline one dimensional axis can be

373 approximated more easily and the ML estimates of dispersal using the 1D and 2D models were
374 similar, further analysis of dispersal distributions via simulations focused on the 1D model.

375 The simulated dispersal distributions are presented in Table 2. All distributions compatible
376 with the empirical isolation by distance regression involved mean dispersal distances greater than
377 123 km. Examination of cumulated distributions reveals that 10% of dispersal events occurred at
378 distances greater than 326 km in all distributions generated and, on average, at distances greater
379 than 1,809 km (Table 3).

380 The ML estimates of the standard deviation of the parent-offspring dispersal distance
381 remained large when a low average mutation rate (10^{-4}) was considered with a point estimate at
382 349 km (95% CI: 114-1,126). The estimate using a high mutation rate scenario (average 10^{-3})
383 yielded substantially larger values for sigma (point estimate 1,103, 95% CI: 361-3,559).
384 Simulated dispersal distributions accounting for the two mutation rates in IBDSIM all yielded an
385 estimate of σ greater than 123 km ($\mu = 10^{-4}$) or 141 km ($\mu = 10^{-3}$) when the census density was
386 used in calculations, or 231 ($\mu = 10^{-4}$) and 259 ($\mu = 10^{-3}$) when the estimate of effective density
387 was used (Appendix 6).

388 Exclusion analysis in GENECLASS detected three putative migrants. The estimate of the 1D
389 isolation-by-distance slope obtained after excluding those 3 individuals was 4.08×10^{-8} (lower
390 and upper bounds of the slope -5.17×10^{-7} and 4.66×10^{-7}) and corresponded to sigma values of
391 795 (95% CI: 236- $+\infty$) or 187 (95% CI: 55- $+\infty$) when considering effective and census density
392 respectively.

393

394 Discussion

395 Allele frequencies at the 17 microsatellites were homogeneous across the sampled area as
396 indicated by the very low estimates of F_{ST} . Only two pairwise exact tests comparing the
397 southwest Florida sample to the southeast Florida and east Texas/Louisiana samples respectively
398 were significant. These three geographic samples did not differ significantly in allele frequencies
399 from any other regional samples, leading to the interpretation that the marginal difference
400 between these localities did not correspond to true barriers to gene flow. This finding was

401 confirmed by the outcome of Bayesian clustering using a spatially explicit approach in TESS
402 which converged toward a single unit and no discontinuity.

403 The lack of divergence among regional samples is consistent with a preliminary assessment
404 based on mitochondrial DNA conducted by Antoni *et al.* (2011). Genetic discontinuities within
405 the sampled area have been evidenced in a variety of other marine and coastal species, in
406 particular between the Gulf of Mexico and the U.S. east coast (Avice 1992), or between
407 populations east and west of Mobile Bay (Karlsson *et al.* 2009; Portnoy & Gold 2012). These
408 reported genetic breaks involved species occupying coastal or estuarine habitats, or species using
409 offshore habitats but displaying characteristics prone to maintaining geographic structure such as
410 limited dispersal abilities. In contrast, species occupying outer shelf habitats similar to those used
411 by the gray triggerfish and dispersing pelagic larvae did not display clear genetic discontinuities
412 across the same geographic area (e.g. red porgy, *Pagrus pagrus*, Ball *et al.* 2007, or the red
413 snapper, *Lutjanus campechanus*, Saillant *et al.* 2010; Hollenbeck *et al.* 2015).

414 The spatial scale of demographic connectivity in gray triggerfish was explored by estimating
415 the parameters of the isolation-by-distance model. Both the moment estimator of Watts *et al.*
416 (2007) and the maximum likelihood estimate in MIGRAINE (Rousset & Leblois 2007, 2012)
417 yielded large estimates of neighborhood sizes with estimates of the parameter σ approaching
418 800 km. Simulated distributions of dispersal distances using different families of functions and
419 different mutation rates yielded average dispersal distances between 123 and 1,323 km.
420 Moreover, examination of the simulated distributions of dispersal distances indicated that 10% of
421 dispersal events resulted in migrations across very long distances from origin (the average 90%
422 percentile was 1,809 km). Interestingly, the relatively high frequency of long distance dispersal
423 events (90% percentile in the hundreds of kilometers) was observed in all simulations, including
424 those where the census population size (which can be considered as an upper bound of effective
425 density) was used, which indicates that the inference that demographic connectivity occurs across
426 long distances is not affected by uncertainties on the value of effective population density. A
427 fraction of immigrants of 10% is usually considered as a threshold below which connected
428 populations are transitioning from demographic dependence to independence (Hastings 1993;
429 Waples & Gaggiotti 2006). While gene flow cannot be easily quantified in terms of a percentage of
430 immigrants in the case of isolation-by-distance, the long distances traveled by a substantial

431 fraction of gray triggerfish before recruiting to benthic habitats and subsequently to breeding
432 populations is consistent with a large degree of demographic dependency of local recruitment
433 from non-local spawning stocks, including those located several hundreds of km from a given
434 recipient benthic habitat. This result contrasts with finding in studies of the demographic
435 connectivity of various reef fishes (e.g. Roberts 1997; Cowen *et al.* 2006; Puebla *et al.* 2012) that
436 concluded that dispersal of ecological significance was occurring within short distances (less than
437 100 km in most cases). The species considered in these studies dispersed larvae over a period
438 limited to a few weeks and usually less than 40 days while gray triggerfish larvae and juveniles
439 remain in the *Sargassum* habitat for 4 to 7 months (Simmons 2008). Thus, although local
440 spawners could contribute to recruitment in the same region if larvae are caught in local eddies
441 (NMFS 2006), the present results indicate that such local retention, if it occurs, is limited and
442 local recruitment is dependent for a large part on the output of spawning populations located at
443 long distances from recipient habitats. An important consequence for management of gray
444 triggerfish populations is that recruitment cannot be predicted from local spawning biomass since
445 it depends for a large part on the output of non-local spawning populations. Instead, recruitment
446 indices may need to be based on the abundance of newly settled juveniles in order to maintain
447 healthy local populations.

448 Inferences based on the isolation-by-distance relationship imply that dispersal was
449 symmetrical along a one dimensional axis. Information on the movement and dynamics of
450 *Sargassum* patches used by gray triggerfish larvae and juveniles is still limited. The peak of the
451 gray triggerfish spawning season occurs in June and July (Simmons and Szedlmayer 2011).
452 During these months *Sargassum* is found in abundance in the Gulf of Mexico and tends to move
453 off the Florida coast and along the Gulf Stream in September (Gower & King 2008). This could
454 favor asymmetric dispersal rates from the Gulf to the Atlantic, a hypothesis that cannot be
455 formally tested using currently available methods to analyze isolation-by-distance. Improved data
456 on the accumulation and movement of *Sargassum* would also be helpful in order to develop more
457 accurate dispersal models for gray triggerfish in the region. Another limitation of the 1D model
458 used in this study is that shorter dispersal routes across open water were not accounted for leading
459 to potential bias during inference of long distance dispersal events in particular in the Gulf of
460 Mexico. Considering dispersal across sections of the open Gulf (e.g. from south Texas to West

461 Florida) in a 2D framework is challenging because gray triggerfish larvae cannot settle in the
462 middle of the Gulf thus violating assumptions of the model. Estimating density is also
463 challenging because the adult habitat is limited to the shelf. An upper bound of density was
464 obtained considering a 20 km wide strip surrounding the 1D lattice and led to estimates of σ
465 consistent with those of the 1D model yet likely under-estimating σ . Thus, while further
466 developments of isolation-by-distance models to allow accounting for the specific characteristics
467 of habitats used by gray triggerfish and the dispersal process would be needed, the inference of
468 large neighborhood sizes and long distance dispersal seems supported under the two models.
469 Another underlying assumption made during inferences on connectivity based on population
470 genetics models is that the population has reached an equilibrium situation. While this cannot be
471 determined easily, repeated temporal sampling could be conducted to confirm the temporal
472 stability of patterns described in this study.

473 The analysis conducted in this work also implicitly neglected the effects of immigration from
474 geographic populations in other portions of the species' range. Gray triggerfish are reported in
475 Central and South America, in Europe and the Mediterranean Sea, and in western Africa (Robins
476 & Ray 1986; Sazonov & Galaktionova 1987). Migrations of gray triggerfish from populations
477 located in the east Atlantic or South America are unlikely considering the long distances involved
478 and large sections of unsuitable habitats for adults in the open Atlantic; Caribbean habitats are
479 closer to the Gulf but the species appears extremely rare in that region (L. Antoni and E. Saillant
480 Unpublished results). However, the impact of rare migrants from divergent populations on
481 estimates of isolation by distance parameters cannot be excluded and was evaluated by omitting
482 possible migrants identified in an exclusion analysis in GENECLASS. The parameters obtained
483 were very similar to those generated using the entire dataset suggesting that estimates are robust
484 to this departure of the 1D model.

485 Gray triggerfish are also present in the southern Gulf of Mexico (e.g. the Bay of Campeche).
486 Populations from the southern Gulf would be expected to be connected to the studied populations
487 and follow the isolation-by-distance pattern described in this study with the additional implication
488 that the effective density estimate would be lower depending on the geographic extent of gray
489 triggerfish south of Texas and the limitations of the 1D model discussed above. Genetic

490 characterization of gray triggerfish in the southern Gulf and study of their abundance is warranted
491 to evaluate this hypothesis and refine current estimates of dispersal parameters.

492 The ratio of effective to census population density was approximately 5.5×10^{-2} . This value is
493 intermediate between the extremely low ratios of effective to census population size (10^{-3} to 10^{-5})
494 reported in studies of some other marine fishes (Turner *et al.* 2002; Hauser *et al.* 2002; Saillant &
495 Gold 2006) and the range (> 0.1) expected in most situations based on demographic models
496 (Nunney & Elam 1994). Estimating effective population size/density is particularly challenging
497 in marine species structured in large connected populations as is the case for gray triggerfish
498 (Hare *et al.* 2011). Methods based on coalescent simulations such as the model used in the
499 present study tend to estimate the size of the overall metapopulation that includes all demes
500 connected to one another by migrations as long as migration is not too low (Hare *et al.* 2011).
501 These methods also integrate the various historical events experienced by the metapopulation
502 over time meaning that it is difficult to determine an appropriate census number that can be
503 matched with the obtained estimates of N_e . The model used in the present study accounted for
504 historical population growth rate of gray triggerfish and thus the estimate of N generated is
505 expected to reflect current/recent N_e , after the detected recent change in population size event
506 (Leblois *et al.* 2014). Very recent changes in population size might not be reflected in the
507 coalescent estimate and the ratio D_e/D_c may be biased if the estimates of census and effective size
508 respectively correspond to different time periods. Alternative methods to estimate
509 contemporaneous effective size such as the linkage disequilibrium (Waples 2006; Waples & Do
510 2010) would have been preferable to match directly census and effective numbers for the same
511 cohorts (Hare *et al.* 2011) but these methods are very imprecise when N_e is greater than 1,000 as
512 was found in the present study. When there is isolation-by-distance, estimates of N_e by the
513 linkage disequilibrium based on samples collected within a breeding window tend to reflect the
514 neighborhood size (Neel *et al.* 2013). This suggests that, even though results from the linkage
515 disequilibrium method lack of precision in the present case, the infinite or very large estimates
516 are consistent with the very large neighborhood size inferred during isolation-by-distance
517 analysis. The census density estimate was derived based on catch data available from the NOAA
518 Office of Science and Technology database for the period that matched genetic sampling and
519 approximates the density of adults present on benthic habitats. This value can be considered an

520 upper bound for population density as it was uncorrected for potential factors likely to lower N_e
521 such as biased sex ratio and variance in reproductive success (Nunney & Elam 1994).

522

523 **Conclusions and management implications**

524 This study used a genetic approach to estimate demographic connectivity among geographic
525 populations of the gray triggerfish, an adult-sedentary reef fish with extended pelagic dispersal.
526 Estimates of the dispersal parameters in an isolation-by-distance framework were consistent with
527 large neighborhoods and dispersal events spread-out over long sections of the shelf habitat used
528 by the species. These estimates suggest a reduced role of local retention in determining local
529 recruitment and a high dependency on the reproductive output of non-local spawning stocks
530 potentially located hundreds or even thousands of kilometers away from recipient benthic
531 habitats. This result contrasts with findings in other reef fishes that disperse pelagic larvae over
532 shorter periods (and distances) and suggests that the longer dispersal in this species is associated
533 with a reduced importance of local retention. Implications for management of populations are
534 significant in that fisheries harvest cannot be managed under the assumption that local biomasses
535 are the major determinant of recruitment. Divergence among geographic regions is insufficient to
536 implement classical mixed stock fisheries models in this case and alternative approaches would
537 need to be developed. Further information on dispersal distribution would be useful in order to
538 better characterize demographic connectivity and develop appropriate models for management of
539 regional fisheries. Studies of the dynamics of formation and movement of *Sargassum* patches in
540 particular will be useful to develop more accurate models predicting dispersal and could then be
541 used to study dispersal in other species that utilize this habitat at early life stages.
542 Contemporaneous estimates of effective population density would also be needed but it will be
543 challenging to generate those estimates based on genetic data. Improved data on life history traits
544 of gray triggerfish would be useful to estimate population size using demographic methods.

545

546 **Acknowledgments**

547 The authors would like to thank B. Walling, D. Devries, and B. Barnett (National Marine
548 Fisheries Service, Panama City laboratory), K. Johnson, A. Debose, and M. Hendon and the
549 scientific staff of the National Marine Fisheries Service Mississippi Laboratory for assistance

550 with collections in the northcentral and northwestern Gulf of Mexico, R. Chapman and M.
551 Reichert for assistance with collections from South Carolina region, Capt. Bill Taylor and staff of
552 the Black Dog Charters for assistance with collection of samples from Jupiter area. Jim Franks
553 provided precious help with coordination of sampling. We also thank Dr. R. Leblois for
554 assistance with running the software MIGRAINE. Funding for this work was provided by the
555 NOAA MARFIN project #NA09NMF4330150.

556

557 **References**

- 558 Almany GR, Berumen ML, Thorrold SR, Planes S, Jones GP (2007) Local replenishment of coral
559 reef fish populations in a marine reserve. *Science*, **316**, 742–744.
- 560 Antoni L, Emerick N, Saillant E (2011) Genetic Variation of Gray Triggerfish in U.S. Waters of
561 the Gulf of Mexico and Western Atlantic Ocean as Inferred from Mitochondrial DNA
562 Sequences. *North American Journal of Fisheries Management*, **31**, 714-721.
- 563 Antoni L, Saillant E (2012) Development and characterization of microsatellite markers in the
564 gray triggerfish (*Balistes capriscus*). *Conservation Genetics Resources*, **4**, 629-631.
- 565 Avise JC (1992) Molecular population structure and the biogeographic history of a regional
566 fauna: a case history with lessons for conservation biology. *Oikos*, **63**, 62-76.
- 567 Avise JC (1998) Conservation genetics in the marine realm. *Journal of Heredity*, **89**, 377-382.
- 568 Ball AO, Beal MG, Chapman RW (2007) Population structure of red porgy, *Pagrus pagrus*, in
569 the Atlantic Ocean. *Marine Biology*, **150**, 1321-1332.
- 570 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful
571 approach to multiple testing. *Journal of the Royal Statistical Society B*, **57**, 289-300.
- 572 Chen C, Durand E, Forbes F, François O (2007) Bayesian Clustering Algorithms Ascertain
573 Spatial Population Structure: A New Computer Program and a Comparison Study. *Molecular*
574 *Ecology Notes*, **7**, 747-756.

- 575 Christie MR, Stallings CD, Johnson DW, Hixon MA (2010) Self-recruitment and sweepstakes
576 reproduction amid extensive gene flow in a coral-reef fish. *Molecular Ecology*, **19**, 1042-1057.
- 577 Cowen RK, Paris CB, Srinivasan A (2006) Scaling of Connectivity in Marine Populations.
578 *Science*, **311**, 522-527.
- 579 Durand E, Chen C, François O (2009a) Tess version 2.1—reference manual. Available from
580 <http://membrestimc.imag.fr/Olivier.Francois/tess.html>. Accessed July 2013.
- 581 Durand E, Jay F, Gaggiotti OE, François O (2009b) Spatial inference of admixture proportions
582 and secondary contact zones. *Molecular Biology and Evolution*, **26**, 1963-1973.
- 583 El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among
584 populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical*
585 *and Applied Genetics*, **92**, 832-839.
- 586 Estoup A, Angers B (1998) Microsatellites and minisatellites for molecular ecology: theoretical
587 and empirical considerations. In: *Advances in molecular ecology* (ed Carvalho G), pp. 55-86.
588 NATO ASI series. IOS Press, Amsterdam, Netherland.
- 589 Franks JS, Hoffmayer ER, Comyns BH, Hendon JR, Blake EM, Gibson DP (2007) Investigations
590 of fishes that utilize pelagic *Sargassum* and frontal zone habitats in Mississippi marine waters
591 and the north central Gulf of Mexico. Final Report presented to the Mississippi Department of
592 Marine Resources, Biloxi, Mississippi, and the U.S. Fish and Wildlife Service, Atlanta,
593 Georgia. Available
594 from <http://gcr1.usm.edu/sargassum/docs/Franks.et.al.%202007.Sargassum.Report.pdf>.
595 Accessed 08 January 2016.
- 596 Goudet J (1995) FSTAT version 1.2: a computer program to calculate *F*-statistics. *Journal of*
597 *Heredity*, **86**(6), 485-486. Available from <http://www2.unil.ch/popgen/softwares/fstat.htm>.
- 598 Gower J, King S (2008) Satellite Images Show the Movement of Floating *Sargassum* in the Gulf
599 of Mexico and Atlantic Ocean. *Nature Precedings*, available
600 from <http://hdl.handle.net/10101/npre.2008.1894.1>. Accessed 08 January 2016.

- 601 Hare MP, Nunney L, Schwartz MK, Ruzzante DE, Burford M, Waples RS, Ruegg K, Palstra F
602 (2011) Understanding and estimating effective population size for practical application in
603 marine species management. *Conservation Biology*, **25**, 438-449.
- 604 Harmelin-Vivien ML, Quéro J-C (1990). Balistidae. p. 1055-1060. In: *Check-list of the fishes of*
605 *the eastern tropical Atlantic (CLOFETA)* (eds. Quéro J-C, Hureau JC, Karrer C, Post A
606 Saldanha S). JNICT, Lisbon; SEI, Paris; and UNESCO, Paris. Vol. 2.
- 607 Hastings A (1993) Complex interactions between dispersal and dynamics: lessons from coupled
608 logistic equations. *Ecology*, **74**, 1362-1372.
- 609 Hauser L, Adcock GJ, Smith PJ, Ramirez JHB, Carvalho GR (2002) Loss of microsatellite
610 diversity and low effective population size in an overexploited population of New Zealand
611 snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences, USA*, **99**, 11742-
612 11747.
- 613 Hollenbeck CM, Portnoy DS, Saillant E, Gold JR (2015) Population structure of red snapper
614 (*Lutjanus campechanus*) in U.S. waters of the western Atlantic Ocean and the northeastern Gulf
615 of Mexico. *Fisheries Research*, **172**, 17-25.
- 616 Hood GM (2010) PopTools version 3.2.5. Available from <http://www.poptools.org>. Accessed 09
617 January 2016.
- 618 Hood PB, Johnson AK (1997) A study of the age structure, growth, maturity schedules and
619 fecundity of gray triggerfish (*Balistes capriscus*), red porgy (*Pagrus pagrus*), and vermilion
620 snapper (*Rhomboplites aurorubens*) from the eastern Gulf of Mexico. MARFIN Final Report.
621 St. Petersburg, Florida, USA.
- 622 Ingram GW Jr. (2001) *Stock structure of gray triggerfish, Balistes capriscus, on multiple spatial*
623 *scales in the Gulf of Mexico*. PhD Dissertation. University of South Alabama, Mobile,
624 Alabama, USA.

- 625 Jing L, Zapfe G, Shao K-T, Leis JL, Matsuura K, Hardy G, Liu M, Tyler J (2015) *Balistes*
626 *capriscus*. In: IUCN 2015. IUCN Red List of Threatened Species. Version 2015-
627 4. <http://www.iucnredlist.org>. Accessed 08 January 2016.
- 628 Johnson, DR, Perry HM, Lyczkowski-Shultz J, Hanisko D. (2009). Red Snapper Larval Transport
629 in the Northern Gulf of Mexico. *Transactions of the American Fisheries Society*, **138**, 458-470.
- 630 Jones G, Almany G, Russ G, Sale PF, Steneck RS, van Oppen MJH, Willis BL (2009) Larval
631 retention and connectivity among populations of corals and reef fishes: history, advances and
632 challenges. *Coral Reefs*, **28**, 307-325.
- 633 Karlsson S, Saillant E, and Gold JR (2009) Population structure and genetic variation of lane
634 snapper (*Lutjanus synagris*) in the northern Gulf of Mexico. *Marine Biology*, **156**, 1841-1855.
- 635 Kritzer JP, Sale PF (2004) Metapopulation ecology in the sea: from Levins' model to marine
636 ecology and fisheries science. *Fish and Fisheries*, **5**, 131-140.
- 637 Leblois R, Estoup A, Rousset F (2009) IBDSim: a computer program to simulate genotypic data
638 under isolation by distance. *Molecular Ecology Resources*, **9**, 107-109.
- 639 Leblois R, Rousset F, Estoup A (2004) Influence of spatial and temporal heterogeneities on the
640 estimation of demographic parameters in a continuous population using individual
641 microsatellite data. *Genetics*, **166**, 1081-1092.
- 642 Leblois R, Pudlo P, Néron J, Bertaux F, Beeravolu CR, Vitalis R, Rousset F (2014) Maximum-
643 Likelihood Inference of Population Size Contractions from Microsatellite Data. *Molecular*
644 *Biology and Evolution*, **31**, 2805-2823.
- 645 Leis JM, McCormick MI (2002) The biology, behavior, and ecology of the pelagic larval stage of
646 coral reef fishes. In: *Coral Reef Fishes: Dynamics and Diversity in a Complex Ecosystem* (ed
647 Sale PF), pp. 171-199. Academic Press, San Diego, California, USA.
- 648 Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity?
649 *Molecular Ecology*, **19**, 3038-3051.

- 650 Neel MC, McKelvey K, Ryman N, Lloyd MW, Short Bull R, Allendorf FW, Schwartz MK,
651 Waples RS (2013) Estimation of effective population size in continuously distributed
652 populations: there goes the neighbourhood. *Heredity*, **111**, 189-199.
- 653 Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, New
654 York USA.
- 655 Nichols S (2004) Derivation of red snapper time series from SEAMAP and groundfish trawl
656 surveys. SEDAR7-DW-01. NMFS, Pascagoula, Mississippi, 28 pp.
- 657 NMFS (2006) Stock Assessment Report of SEDAR 9: Gulf of Mexico Gray Triggerfish.
658 Charleston, South Carolina. 195 pp.
- 659 NMFS (2011) SEDAR 9 Update Stock Assessment Report Gulf of Mexico Gray Triggerfish.
660 Tampa, Florida, USA. 270 pp.
- 661 Nunney L, Elam DR (1994) Estimating the effective population size of conserved populations.
662 *Conservation Biology*, **8**, 175-184.
- 663 Paetkau D, Slade R, Burden M, Estoup A (2004) Direct, real-time estimation of migration rate
664 using assignment methods: a simulation-based exploration of accuracy and power. *Molecular*
665 *Ecology*, **13**, 55-65.
- 666 Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: A
667 Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of*
668 *Heredity*, **95**, 536-539.
- 669 Portnoy, DS, Gold JR (2012) Evidence of multiple vicariance in a marine suture-zone in the Gulf
670 of Mexico. *Journal of Biogeography*, **39**, 1499-1507.
- 671 Puebla O, Bermingham E, Guichard F (2009) Estimating dispersal from genetic isolation by
672 distance in a coral reef fish (*Hypoplectrus puella*). *Ecology*, **90**, 3087-3098.
- 673 Puebla O, Bermingham E, McMillan O (2012) On the spatial scale of dispersal in coral reef
674 fishes. *Molecular Ecology*, **21**, 5675-5688.

- 675 Rannala B, Mountain JL (1997) Detecting immigration by using multilocus
676 genotypes. *Proceedings of the National Academy of Sciences USA*, **94**, 9197-9201.
- 677 Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact
678 tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.
- 679 Roberts CM (1997) Connectivity and management of Caribbean coral reefs. *Science*, **278**, 1454-
680 1457.
- 681 Robins, CR, Ray GC (1986) *A field guide to Atlantic coast fishes of North America*. (ed Peterson
682 RT). Houghton Mifflin Harcourt, Boston, Massachusetts, USA.
- 683 Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under
684 isolation by distance. *Genetics*, **145**, 1219-1228.
- 685 Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*.
686 **13**, 58-62.
- 687 Rousset F (2008a) Genepop'007: a complete reimplement of the Genepop software for
688 Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.
- 689 Rousset F (2008b) Dispersal estimation: Demystifying Moran's I. *Heredity*, **100**, 231-232.
- 690 Rousset F, Leblois R (2007) Likelihood and approximate likelihood analyses of genetic structure
691 in a linear habitat: performance and robustness to model mis-specification. *Molecular Biology
692 and Evolution*, **24**, 2730-2745.
- 693 Rousset F, Leblois R (2012) Likelihood-based inferences under a coalescent model of isolation
694 by distance: two-dimensional habitats and confidence intervals. *Molecular Biology and
695 Evolution*, **29**, 957-973.
- 696 Saillant E, Bradfield SC, Gold JR (2010) Genetic variation and spatial autocorrelation among
697 young-of-the-year red snapper (*Lutjanus campechanus*) in the northern Gulf of Mexico. *ICES
698 Journal of Marine Science*, **67**, 1240-1250.

- 699 Saillant E, Gold JR (2006) Population structure and variance effective size of red snapper
700 (*Lutjanus campechanus*) in the northern Gulf of Mexico. *Fishery Bulletin*, **104**, 136–148.
- 701 Sambrook JE, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn.
702 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.
- 703 Sazonov, YG, Galaktionova AI (1987) Some data on the morphometrics of the grey triggerfish,
704 *Balistes carolinensis*, of the central-eastern Atlantic. *Journal of Ichthyology*, **27**, 173-176.
- 705 Shanks AL (2009) Pelagic Larval Duration and Dispersal Distance Revisited. *Biological Bulletin*,
706 **216**, 373-385.
- 707 Simmons CM (2008) Gray triggerfish, *Balistes capriscus*, reproductive behavior, early life
708 history, and competitive interactions between red snapper, *Lutjanus campechanus*, in the
709 northern Gulf of Mexico. PhD Dissertation. Auburn University, Auburn, Alabama, USA.
- 710 Simmons CM, Szedlmeyer ST (2011) Recruitment of Age-0 Gray Triggerfish to Benthic
711 Structured Habitat in the Northern Gulf of Mexico. *Transactions of the American Fisheries*
712 *Society*, **140**, 14-20.
- 713 Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model
714 complexity and fit (with discussion). *Journal of the Royal Statistical Society B*, **64**, 191-232.
- 715 Thorrold SR, Jones GP, Hellberg ME, Burton RS, Swearer SE, Neigel JE, Morgan SG, Warner
716 RR (2002) Quantifying larval retention and connectivity in marine populations with artificial
717 and natural markers. *Bulletin of Marine Science*, **70**, 291-308.
- 718 Turner TF, Wares JP, Gold JR (2002) Genetic effective size is three orders of magnitude smaller
719 than adult census size in an abundant, estua-rine-dependent marine fish (*Sciaenops ocellatus*).
720 *Genetics*, **162**, 1329-1339.
- 721 Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software
722 for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology*
723 *Notes*, **4**, 535-538.

- 724 Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high
725 gene flow species. *Journal of Heredity*, **89**, 438-450.
- 726 Waples RS (2006) A bias correction for estimates of effective population size based on linkage
727 disequilibrium at unlinked gene loci. *Conservation Genetics*, **7**, 167-184.
- 728 Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly
729 variable genetic markers: a largely untapped resource for applied conservation and evolution.
730 *Evolutionary Applications*, **3**, 244-262.
- 731 Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic
732 methods for identifying the number of gene pools and their degree of connectivity. *Molecular*
733 *Ecology*, **15**, 1419-1439.
- 734 Watts PC, Rousset F, Saccheri IJ, Leblois R, Kemp SJ, Thompson DJ (2007) Compatible genetic
735 and ecological estimates of dispersal rates in insect (Coenagrion mercuriale: Odonata:
736 Zygoptera) populations: analysis of 'neighbourhood size' using a more precise estimator.
737 *Molecular Ecology*, **16**, 737-751.
- 738 Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure.
739 *Evolution*, **38**, 1358-1370.
- 740 Wells RJD, Rooker JR (2004) Spatial and temporal patterns of habitat use by fishes associated
741 with *Sargassum* mats in the northwestern Gulf of Mexico. *Bulletin of Marine Science*, **74**, 81-
742 99.
- 743 Wilson CA, Nieland DL, Stanley AL (1995) Age, growth and reproductive biology of gray
744 triggerfish (*Balistes capriscus*) from the northern Gulf of Mexico commercial harvest. Final
745 Report. Coastal Fisheries Institute, Louisiana State University, Baton Rouge, Louisiana, USA.

746

747 **Data accessibility**

748 Genotypes of all individual samples at 17 microsatellites: uploaded on the Aquila repository of
749 the University of southern Mississippi as part of the gray triggerfish genetics project.

750 DOI: <https://doi.org/10.18785/gtg.ds.01>

751

752 **Author contributions**

753 E. Saillant and L. Antoni designed the study, E. Saillant coordinated sample acquisition, L.

754 Antoni performed data acquisition, E. Saillant and L. Antoni analyzed the data and wrote the

755 manuscript.

Author Manuscript

Table 1 Estimates of F_{ST} (Weir and Cockerham θ) (upper diagonal) and probability that $F_{ST} = 0$ (lower diagonal) for pair-wise comparisons of microsatellite allele distributions between gray triggerfish geographic samples. Probability values that differed significantly from zero following correction for multiple tests are in bold

	STX	ETX-LA	MS-WF	SWF	SEF	SC
STX		0.0007	0.0003	0.0004	0.0018	-0.0002
ETX-LA	0.182		0	0.0008	0.0006	-0.0006
MS-WF	0.543	0.048		0.0004	0.0012	-0.0006
SWF	0.163	0.018	0.259		0.0018	0.0003
SEF	0.220	0.203	0.297	0.003		0.0015
SC	0.514	0.618	0.323	0.265	0.098	

Author Manuscript

Table 2 Parameters of simulated distributions yielding isolation-by-distance slopes comparable to that of the empirical dataset (point estimate and upper bound). D: population density; μ_x : mean (simulated) dispersal distance; σ : standard deviation of parental position relative to offspring position; sim.: simulated; est.: estimated; P-value: Range of Mantel test P-values (10 000 permutations) in the 3 replicates

Model	μ	σ (sim./est.)	IBD Slope	95%-	95%+	P-value
D _c = 10 ind./km						
Empirical		914	3.10E-08	-5.24E-07	4.61E-07	0.159
Pareto (M=0.995 ; n=1.16)	482	1,509/293	3.02E-07	1.25E-07	4.99E-07	0.0000- 0.0002
Pareto (M=0.97 ; n=0.92)	1323	2,666/938	2.94E-08	-9.93E-08	1.88E-07	0.0248- 0.0670
Geometric (m=0.98 ; g=0.993)	140	203/242	4.41E-07	2.50E-07	6.48E-07	0.0000
Geometric (m=0.95 ; g=0.999)	950	1,400/1,031	2.43E-08	-1.02E-07	1.74E-07	0.0000
Sichel ($\gamma=-0.0005$; $\xi=15000$; $\Omega=0.002$)	420	1,037/318	2.56E-07	4.89E-08	6.64E-07	0.0000- 0.0004
Sichel ($\gamma=-0.002$; $\xi=15000$; $\Omega=0.001$)	505	1,263/806	3.98E-08	-8.92E-08	2.05E-07	0.0113- 0.0794
D _c = 175 ind./km						

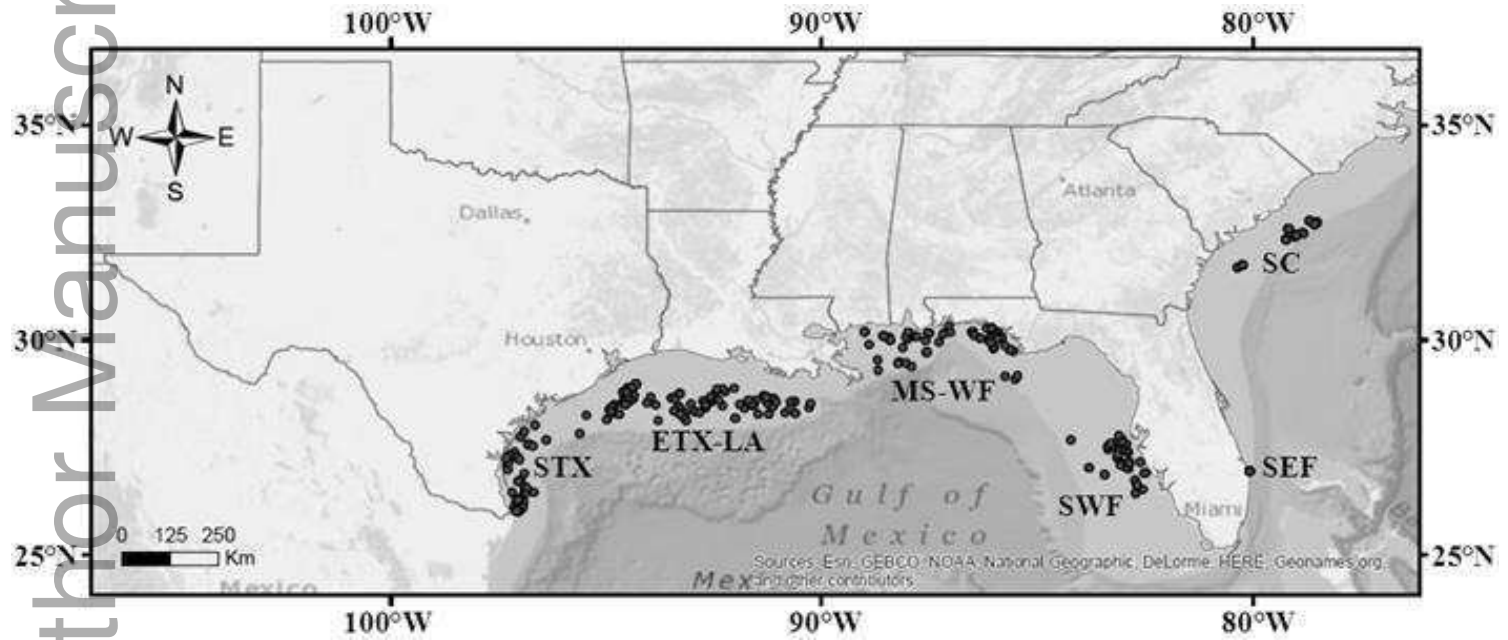
Empirical		215	3.10E-08	-5.24E-07	4.61E-07	0.159
Pareto (M = 0.95; n = 0.98)	1047	2,300/219	2.99E-08	-8.65E-08	1.51E-07	0.0143- 0.1863
Geometric (m=0.98 ; g=0.992)	123	175/213	3.14E-08	-6.28E-08	1.45E-07	0.0085- 0.0522
Sichel ($\gamma=-0.001$; $\xi=10000$; $\Omega=0.004$)	278	660/212	3.19E-08	-9.57E-08	1.69E-07	0.0011- 0.1525

Author Manuscript

Table 3 Percentile distribution of the simulated functions compatible with the isolation-by-distance slope estimated during the study

Distribution parameter	Percentile			
	25	50	75	90
Pareto (M=0.995 ; n=1.16)	3	12	137	1,247
Pareto (M=0.97 ; n=0.92)	12	166	1,519	4,866
Geometric (m=0.98 ; g=0.993)	40	97	196	326
Geometric (m=0.95 ; g=0.999)	238	643	1,336	2,251
Sichel ($\gamma=-0.0005$; $\xi=15,000$; $\Omega=0.002$)	14	67	356	1,205
Sichel ($\gamma=-0.002$; $\xi=15,000$; $\Omega=0.001$)	12	65	403	1,485
Pareto (M = 0.95; n = 0.98)	6	75	912	3,889
Geometric (m=0.98 ; g=0.992)	35	85	172	286
Sichel ($\gamma=-0.001$; $\xi=10,000$; $\Omega=0.004$)	13	54	241	726

Author Manuscript



Author Manuscript

