

## Exploring the structure of genetic variation and the influences of demography on effective population size in the gag grouper *Mycteroperca microlepis* (Goode & Bean)

N. K. JUE

*Department of Biological Science, Florida State University, Tallahassee,  
FL 32306–1100, U.S.A.*

Using 11 microsatellite markers, genetic analyses of three successive year-classes of gag *Mycteroperca microlepis* juveniles across the north-eastern Gulf of Mexico revealed a lack of spatial structure and very little temporal variation between year-classes. These results are consistent with long-term effective population sizes on the order of 30 000 adults. The importance of reproductive-style and sex-ratio variation is discussed as an important influence on long-term effective sizes.

© 2006 The Author

Journal compilation © 2006 The Fisheries Society of the British Isles

Key words: effective population size; *Mycteroperca microlepis*; sex ratio.

As many marine species are heavily exploited due to fisheries selective pressures, which could have adverse influences on standing genetic variation in populations (Olsen *et al.*, 2004), it is worthwhile to attempt to understand the factors that influence reproductive success and determine the effective population size in fisheries systems. The status of gag *Mycteroperca microlepis* (Goode & Bean) in the Gulf of Mexico raises a number of questions about whether present demographic conditions and harvesting rates might be affecting the evolutionary future of the stock. Fishing pressure has been identified as skewing sex ratios (male:female) from historical levels of 1:6 to 1:30 or greater (Coleman *et al.*, 1996). In this study, 11 microsatellite marker genes were used to examine spatial and temporal genetic structure of young-of-the-year (YOY) gag in the eastern Gulf, estimate long-term effective population size, and assess some probable effects contributing to the estimated value of effective population size.

Gag, a member of the *Serranidae*, is a protogynous hermaphrodite found throughout the Atlantic coastal areas of North, Central, and South America, where their current centre of abundance appears to be in the north-eastern Gulf of Mexico (Koenig *et al.*, 2000). Gag spawn in offshore aggregations located on the shelf edge (*c.* 80–100 m in depth) with peak spawning activity occurring in late winter and early spring and have some potential for extreme

long-distanced dispersal (*e.g.* hundreds of km) (McGovern *et al.*, 2005), particularly during prespawning life stages, while reproductively active males, and possibly large females, seem to exhibit high site fidelity, staying close to aggregation sites year round (C. Koenig, pers. comm.). Individuals begin to mature sexually as females around the age of 3 or 4 years and may change into males sometime between the ages of 6 and 8 years (McGovern *et al.*, 1998).

Three separate cohorts of YOY gag were sampled from 12 sites along the West Florida Shelf in the years 2003, 2004 and 2005. These sites were spread out over *c.* 900 km of coastline, from St Andrews Bay near Panama City to Cape Romano near Naples (Fig. 1). A small percentage of these individuals were collected through the assistance of commercial bait shrimpers (19 fish) and NOAA Fisheries (28 fish). These additional samples were in close proximity to areas sampled with otter trawls and were included as part of the genetic collections for these sites. When densities were sufficient, between 30 and 50 fish were sampled for tissues at each site. All tissues were stored in 1% *n*-laurylsarcosine, 8 M urea, 20 mM sodium phosphate, and 1 mM EDTA, pH 6.8. Genomic DNA was extracted using magnetic beads methodology (Agencourt, Inc., Beverly, MA, U.S.A.).

Microsatellite primers developed for gag (Chapman *et al.*, 1999), black grouper *Mycteroperca bonaci* (Poey) (Zatcoff *et al.*, 2002), red grouper *Epinephelus morio* (Valenciennes) (Zatcoff *et al.*, 2002) and Hawaiian grouper *Epinephelus quernus* Seale (Rivera *et al.*, 2003), were assayed using fluorescently labelled primers (IDTDNA, Applied Biosystems, Foster City, CA, U.S.A.). Samples were

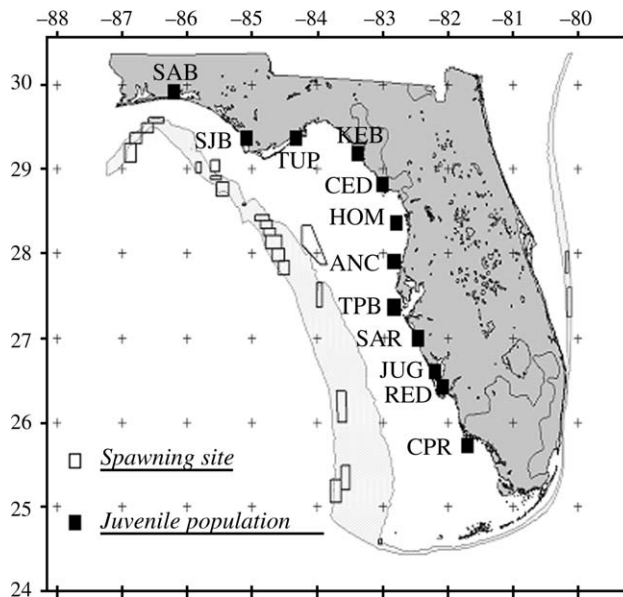


FIG. 1. Map of the West Florida Shelf showing areas of spawning activity (□) and sites sampled for juvenile gag (■). SAB, St Andrews Bay; SJB, St Joe Bay; TUP, Turkey Point; KEB, Keaton Beach; CED, Cedar Key; HOM, Homassassa; ANC, Anclote Key; TPB, Tampa Bay and St Petersburg; SAR, Sarasota; JUG, Jug Creek Shoal; PIS, Pine Island Sound; CPR, Cape Romano.

analysed on an Applied Biosystems (ABI) 3130xl Genetic Analyzer with Capillary Electrophoresis. Data were read and analysed using ABI Genemapper Software version 4.0. After assaying 32 possible primer sets in 50 individuals for positive amplification and polymorphism, 11 loci were found to be polymorphic and easily scoreable. All samples then underwent polymerase chain reaction (PCR) amplification and subsequent genotyping at these 11 loci. Final PCR conditions consisted of concentrations of buffer at 1X, primers at 0.3  $\mu\text{M}$  (forward and reverse), 0.2 mM dNTPs, 2 mM  $\text{Mg}^{2+}$ , 0.05 units  $\mu\text{L}^{-1}$  Invitrogen Taq, and 1  $\mu\text{g } \mu\text{L}^{-1}$  BSA. PCR protocols varied among loci.

Data were examined at the level of site and year. The expected and observed numbers of homozygotes were calculated using MICROCHECKER v. 2.2.1 (van Oosterhout *et al.*, 2003). Conformation of observed genotypes to Hardy–Weinberg equilibrium (HWE) was assessed. Also, probabilities that loci were at genotypic linkage equilibrium in each population were estimated by Markov chain-based approximations of exact tests with the programme FSTAT v. 2.9.3.2 (Goudet, 2001). A sequential Bonferroni correction was applied to account for multiple comparisons (Rice, 1989). Pair-wise  $F_{\text{ST}}$  values were calculated to identify any significant spatial or temporal pattern of differentiation among all sites and probabilities of significant differentiation were assessed using the permutation methods implemented by the programme FSTAT (Goudet, 2001). The significance of these estimates was assessed by 1000 permutations. A sequential Bonferroni correction was applied to account for multiple comparisons (Rice, 1989). A hierarchical analysis, AMOVA, was also applied to the data using Arlequin v. 2.000 (Michalakis & Excoffier, 1996). The significance of these estimates was assessed by 1000 permutations.

To describe long-term effective population size ( $N_e$ ) assuming a Brownian stepwise mutational model, the programme MIGRATE v. 2.1.5 (Beerli & Felsenstein, 2001) was used to estimate the parameter  $\theta$ . This parameter is defined as  $\theta = 4N_e\mu$ , where  $\mu$  is the mutation rate of substitution at a locus or group of loci. Ten replicates of 20 individuals were randomly selected from the complete data (all years grouped together) and each year of samples. Each year was run separately to ensure stability of result. The truncation of the data set was done because the usage of the complete data set proved to be computationally prohibitive. Large numbers of individuals necessitate very long runs of Markov-chain Monte-Carlo methods to fully explore tree parameter space, while little in precision may be gained by increasing the number of individuals (P. Beerli, pers. comm.). A uniform prior distribution over the interval  $\{0, 50.0\}$  was chosen for  $\theta$   $\{0, 50.0\}$  and a Bayesian approach was implemented for simulation. All runs consisted of 10 independent chains and sampled 10 000 000 genealogies with a multiple Markov chain, static heating scheme. Due to multiple modes in Bayesian posterior distributions, scaling of loci mutational rates by number of alleles per locus across all samples was employed and examination of posterior distributions proved stable. To obtain greater stability in the result, final estimation of  $\theta$  and associated credible intervals was done by averaging estimates of multi-locus median scores and credible interval ranges of 10 replicate runs of 40 individuals with 500 'bins' posterior distributions. Median scores should give the most stable and reliable estimate of the true value with this method (Beerli, 2006).

Mutation rates were estimated following a protocol developed by Turner *et al.* (2002). A range of mutation rates for microsatellites was used to assess upper and lower bounds of probable rates. Microsatellites are believed to have mutation rates between  $1 \times 10^{-3}$  and  $1 \times 10^{-5}$  (Jarne & Lagoda, 1996). Since  $N_e$  should be the same across all loci, differences in  $\theta$  among loci should be due to differences in mutation rates. Under this assumption, each sample group can be treated as a replicate estimate of the relative  $\mu$  among loci. To encompass the range of probable  $\mu$ , the locus with the largest values for  $\theta$  was assumed to be  $\mu_{\max} = 1 \times 10^{-3}$ . The other ( $j$ ) loci were then scaled using the ratio of  $\theta_j \theta_{\max}^{-1}$  times the assumed value of  $\mu_{\max}$ . The same process was done using the locus with the lowest value of  $\theta$ , assigning it a  $\mu_{\min} = 1 \times 10^{-5}$ , and using to this ratio of  $\theta_j \theta_{\min}^{-1}$  times the assumed value of  $\mu_{\min}$  to scale the other ( $j$ ) loci. Ranges of effective population size (97.5% credible intervals) were calculated using MIGRATE-generated values of  $\theta$  and the midpoint of the interval of harmonic mean estimates of  $\theta_{\min}$  and  $\theta_{\max}$  scaled mutation rates.

A total of 719 YOY individuals were caught across all 3 years of sampling (208, 279 and 232 for 2003, 2004 and 2005, respectively). All loci were considered to be independent as tests for linkage equilibrium proved to be positive. While a few isolated departures from HWE were noted (seven out of 176 analyses), they were on the scale of what might be expected from chance alone (<5%). There was also no consistent pattern of deviation due to any one locus. Only one locus *GAG 038* registered two incidences of HWE departures at the site level, although when all sites within each year were pooled, *GAG 038* did not deviate from HWE. Therefore, the presence of null alleles was not considered to be a significant issue in this analysis.

There were no differences in allele frequencies among comparison groups at either a temporal or spatial level. Pair-wise comparisons revealed very little differentiation between sites within years and between sites across years (Table I). Two sites (TUP 2003 and SAR 2004) seemed to account for most of those differences. After correcting for multiple comparisons, none of these  $F_{ST}$  values were significant. Pair-wise comparisons between year-classes yielded a similar result with low, insignificant values of  $F_{ST}$ . AMOVA results supported this same finding by attributing very little in variance components to comparisons among years and among sites within years (almost 100% attributed to within sites sources of variation).

Since there was little spatial or temporal genetic differentiation and no strong evidence for deviations from HWE, the breeding population was considered to be effectively panmictic. Theta estimates for the entire sample, pooled across years, did not differ from estimates for each year treated as an independent measure. Final MIGRATE runs yielded  $\theta$  values with a median score of 12.39. Mutation rate was estimated to be  $2 \times 10^{-4}$ , since this value was essentially the midpoint of the range explored, a  $\mu = 1 \times 10^{-4}$  was employed to calculate values of  $N_e$  to limit underestimation biases. Derived values of  $N_e$  were on the order of 30 000 [ $N_e = 30,975$ , 97.5% credible interval {22,750, 38,100}].

Thus, gag in the north-eastern Gulf of Mexico exhibit an estimate of  $N_e$  of *c.* 30 000 and genetic homogeneity among sites and years. Despite presumed site fidelity of adults,  $F$ -statistics revealed very low levels of differentiation

TABLE I.  $F_{ST}$  value for pair-wise comparison of sites (see Fig. 1). Numbers in bold and\* are significant at  $P < 0.05$ . After correction for multiple comparisons, none of these comparisons is considered to be significant. Number of individuals in each sample is shown in parentheses under site heading in top row

	SAB 2003 (39)	SJB 2003 (50)	TUP 2003 (41)	SAR 2003 (47)	PIS 2003 (31)	SAB 2004 (38)	SJB 2004 (50)	TUP 2004 (44)	CED 2004 (48)	TPB 2004 (47)	SAR 2004 (52)	SAB 2005 (49)	SJB 2005 (30)	TUP 2005 (42)	JUG 2005 (58)	PIS 2005 (53)
SAB 2003	0															
SJB 2003	-0.0006	0														
TUP 2003	0.0024	-0.0004	0													
SAR 2003	0.0011	0.0001	0.0026	0												
PIS 2003	-0.0030	0.0024	0.0045	-0.0003	0											
SAB 2004	-0.0029	-0.0026	-0.0007	-0.0003	0.0004	0										
SJB 2004	-0.0003	-0.0014	0.0008	-0.0005	0.0018	-0.0008	0									
TUP 2004	0.0001	-0.0014	0.0018	-0.0014	-0.0004	-0.0006	-0.0014	0								
CED 2004	-0.0004	0.0009	<b>0.0033*</b>	0.0019	-0.0001	-0.0003	-0.0001	-0.0033	0							
TPB 2004	-0.0007	0.0005	<b>0.0049*</b>	0.0005	0.0010	-0.0009	-0.0004	-0.0039	-0.0001	0						
SAR 2004	0.0009	0.0006	0.0039	0.0009	-0.0027	0.0011	<b>0.0035*</b>	0.0008	0.0024	<b>0.0035*</b>	0					
SAB 2005	-0.0015	-0.0010	0.0009	-0.0007	0.0006	-0.0015	-0.0014	-0.0015	0.0017	-0.0035	0.0016	0				
SJB 2005	-0.0027	-0.0029	-0.0020	-0.0042	-0.0012	-0.0054	-0.0024	-0.0031	-0.0035	-0.0019	-0.0005	0.0005	0			
TUP 2005	0.0009	-0.0004	-0.0002	0.0003	0.0003	-0.0025	-0.0018	-0.0033	-0.0014	-0.0002	0.0024	-0.0010	-0.0026	0		
JUG 2005	0.0006	-0.0018	0.0022	0.0015	-0.0001	-0.0027	-0.0005	-0.0005	-0.0005	0.0006	-0.0005	0.0005	-0.0032	-0.0008	0	
PIS 2005	0.0009	0.0005	<b>0.0054*</b>	0.0022*	-0.0021	-0.0023	0.0004	-0.0007	-0.0007	0.0017	0.0024	0.0027	-0.0041	0.0011	-0.0010	0

between both spatial and temporal comparison groups of juvenile fish. Yearly settlement patterns could be the result of one genetically distinct subpopulation dominating all reproduction; however, given the temporal and spatial scales considered in this study, that seems unlikely. Larvae from successfully reproducing spawners appear to be mixed across years and a very large geographical area; the distance sampled in this study is a little less than a third of the entire U.S. Atlantic coast. A previous study by Chapman *et al.* (1999), which used three of the markers implemented in this study, observed some deviations from HWE and postulated that these differences were either a result of high variance in reproductive success, causing the above-described temporal Wahlund effect within samples by pooling multiple settlement events or year-classes, or increased inbreeding in the population due to fishing practices. The absence of deviations in the work reported here may be due to differences in sampling methods (no pooling of year-class at each site), molecular methods (use of sequence-based fragment lengths instead of gel scoring), or genuine differences between the stocks in the two studies.

The reproductive style of sequential hermaphroditism is probably a strong influence on estimates of  $N_e$  for gag. Examination of partial derivatives for the following standard equations for  $N_e$  of organism with separate sexes, such as the following:  $N_e^{-1} = (4N_m)^{-1} + (4N_f)^{-1}$  (Wright, 1931), shows that where proportion male,  $m$ , is  $<0.41$  (c. a 2:5 sex ratio) changes in the number of males,  $N_m$ , have a greater effect on rates of change in  $N_e$  than changes in  $N$ . Thus, shifts in demography relative to its life history may contribute more to the amount of genetic variation maintained in a population than the fluctuation of sheer numbers of individuals. Gag adhere to this condition for  $m$  both historically and currently. Additionally, there could also be some sort of process that is indicative of species that exhibit similar patterns of dispersal or life histories in the Gulf of Mexico as similar types of results have been reported in a co-occurring species, the red snapper *Lutjanus campechanus* (Poey) (Saillant & Gold, 2006). To put this value of  $N_e$  in some context, a rough estimate, conservative with respect to overestimation, of current census size would yield a number on the order of a few million reproductively active individuals ([www.sefsc.noaa.gov](http://www.sefsc.noaa.gov)) and thus an approximation of the  $N_e:N$  ratio in the order of 0.01. The relationship between the long-term estimate reported here and a contemporary estimate of census size maybe problematic due to the different time scales under consideration (Waples, 2005). Ideally, a comparison with an estimate of contemporary  $N_e$  such as a temporal sampling method would be done to further validate results reported here. Currently, temporally sampled data for such a comparison are unavailable, but are being pursued.

I would like to thank J. Travis, F. Coleman and C. Koenig for all their support, council and comment in the writing of this paper, R. Chapman and A. Ball for all their help with molecular methods, D. Levitan for usage of his laboratory, P. Beerli for all the conversation, and three anonymous reviewers and D. Ruzzante for their helpful comments. Additionally, special thanks needs to be extended to the numerous people and organizations who provided field assistance, consultation about sampling sites, access to areas, and, of course, places for me to stay. This research was funded through the National Marine Fisheries Service Co-operative Research Program Grant No. NA04NMF4540213, as well as PADI Project Aware, Florida State University

University Fellowship Program, the National Wildlife Refuge Centennial Scholar Program, and the EPA STAR Fellowship Program. None of the above organizations officially endorses this publication and the views expressed herein may not reflect the views of these organizations.

### References

- Beerli, P. (2006). Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. *Bioinformatics* **22**, 341–345.
- Beerli, P. & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4563–4568.
- Chapman, R. W., Sedberry, G. R., Koenig, C. C. & Eleby, B. M. (1999). Stock identification of gag, *Mycteroperca microlepis*, along the southeast coast of the United States. *Marine Biotechnology* **1**, 137–146.
- Coleman, F. C., Koenig, C. C. & Collins, L. A. (1996). Reproductive styles of shallow-water groupers (Pisces: Serranidae) in the eastern Gulf of Mexico and the consequences of fishing spawning aggregations. *Environmental Biology of Fishes* **47**, 129–141.
- Jarne, P. & Lagoda, P. J. L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution* **11**, 424–429.
- Koenig, C. C., Coleman, F. C., Grimes, C. B., Fitzhugh, G. R., Scanlon, K. M., Gledhill, C. T. & Grace, M. (2000). Protection of fish spawning habitat for the conservation of warm-temperate reef-fish fisheries of shelf-edge reefs of Florida. *Bulletin of Marine Science* **66**, 593–616.
- McGovern, J. C., Wyanski, D. M., Pashuk, O., Manooch, C. S. & Sedberry, G. R. (1998). Changes in the sex ratio and size at maturity of gag, *Mycteroperca microlepis*, from the Atlantic coast of the southeastern United States during 1976–1995. *Fishery Bulletin* **96**, 797–807.
- McGovern, J. C., Sedberry, G. R., Meister, H. S., Westendorff, T. M., Wyanski, D. M. & Harris, P. J. (2005). A tag and recapture study of gag, *Mycteroperca microlepis*, off the southeastern US. *Bulletin of Marine Science* **76**, 47–59.
- Michalakis, Y. & Excoffier, L. (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics* **142**, 1061–1064.
- Olsen, E. M., Heino, M., Lilly, G. R., Morgan, M. J., Brattey, J., Ernande, B. & Dieckmann, U. (2004). Maturation trends indicative of rapid evolution preceded the collapse of northern cod. *Nature* **428**, 932–935.
- Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43**, 223–225.
- Rivera, M. A. J., Graham, G. C. & Roderick, G. K. (2003). Isolation and characterization of nine microsatellite loci from the Hawaiian grouper *Epinephelus quernus* (Serranidae) for population genetic analyses. *Marine Biotechnology* **5**, 126–129.
- Saillant, E. & Gold, J. R. (2006). Population structure and variance effective size of red snapper (*Lutjanus campechanus*) in the northern Gulf of Mexico. *Fishery Bulletin* **104**, 136–148.
- 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.
- Waples, R. S. (2005). Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology* **14**, 3335–3352.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.
- Zatcoff, M. S., Ball, A. O. & Chapman, R. W. (2002). Characterization of polymorphic microsatellite loci from black grouper, *Mycteroperca bonaci* (Teleostei: Serranidae). *Molecular Ecology Notes* **2**, 217–219.

### Electronic References

- Goudet, J. (2001). *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3)* (updated from Goudet, 1995). <http://www.unil.ch/izea/software/fstat.html>
- van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. & Shipley, P. F. (2003). *Microchecker, 2.2.1*. Available at <http://www.microchecker.hull.ac.uk/>