Sibship Reconstruction Demonstrates the Extremely Low Effective Population Size of Striped Bass Morone Saxatilis in the Santee-Cooper System, South Carolina, USA

Jin-Xian Liu

Bert Ely
University of South Carolina - Columbia, ely@sc.edu

Follow this and additional works at: https://scholarcommons.sc.edu/biol_facpub

Part of the Biology Commons

Publication Info
Sibship reconstruction demonstrates the extremely low effective population size of striped bass *Morone saxatilis* in the Santee-Cooper system, South Carolina, USA

Jin-Xian Liu, Bert Ely

Department of Biological Sciences, University of South Carolina, Columbia, SC, 29208, USA

*Key words:* sweepstakes reproductive success, effective population size, sibship reconstruction, striped bass, *Morone saxatilis*

*Corresponding author:*

Dr. Bert Ely

Department of Biological Sciences, University of South Carolina, Columbia, SC, 29208, USA

*Tel.:* 803-777-2768; *Fax:* 803-777-4002; *email:* ely@sc.edu

*Running head:* Large variance in reproductive success
Abstract

For organisms with great fecundity and high mortality in early life stages, such as shellfish or fishes, the need to match reproductive activity with environmental conditions conducive to spawning, fertilization, larval development, and recruitment may result in extreme variance in reproductive success among individuals. The main objective of this study was to investigate evidence of large variance in the reproductive success of the striped bass *Morone saxatilis* in the Santee-Cooper system, South Carolina, USA. Seven microsatellite loci were analyzed in 603 recruits representing three yearly cohorts from 1992 to 1994, and a group analysis was performed to identify full sib families. Large variance in reproductive success was detected, with a few large, full-sib families contributing disproportionately to each of the cohorts. The severity of sweepstakes reproductive success varied among cohorts depending on environmentally imposed mortality. Estimations of the effective number of breeders in these long-lived fish ranged from 24 in 1992 to 44 in 1994. Furthermore, the estimated genetic effective population size (*N*<sub>e</sub> = 93) is roughly four orders of magnitude lower than estimates of adult census size (*N* = 362,000). Furthermore, the presence of large full-sib families indicates that striped bass engage in pair mating in the wild. Heterogeneity in genetic composition was also observed among cohorts, suggesting that genetically different adults contribute to different cohorts and that chance rather than fitness variation determines reproductive success.
Introduction

Effective population size \((N_e)\) is the most important population parameter in evolutionary and conservation biology, not only because \(N_e\) determines levels of genetic diversity, but also because \(N_e\) affects the rate of genetic drift and the effectiveness of natural selection (Wright 1931; Frankham 1996). Because \(N_e\) may not reflect the census population size \((N)\), the ratio \(N_e/N\) can be used as a convenient indicator of the extent of genetic variation expected in a population (Hedrick 2005). Theoretically, the \(N_e/N\) ratio is expected to be \(~0.5\) for a wide range of demographic and reproductive scenarios (Nunney 1993, 1996). Palstra & Ruzzante (2008) reviewed estimates of \(N_e/N\) ratios over a number of species and reported a median value of 0.14. Populations with low \(N_e/N\) ratios may suffer reduced capacity to respond to changing or novel environmental pressures, inbreeding depression, and accumulation of deleterious alleles (Higgins & Lynch 2001; Palstra & Ruzzante 2008), especially when \(N\) is a relatively small number (Frankham 1996; Frankham et al. 2003; Frankham 2005).

Recent studies have documented \(N_e/N\) ratios that are several orders of magnitude lower than 0.5 in a number of species (Hedgecock 1994; Hauser et al. 2002; Turner et al. 2002; Hoarau et al. 2005; Gomez-Uchida & Banks 2006). These species of marine shellfish and fish are generally characterized by high fecundities and very high mortalities in their early life stages (type \(\mu\) survivorship curves). Hedgecock (1994) suggested that, for species with type \(\mu\) survivorship, the bulk of recruited young in a given spawning season may be from a small number of parents because of the necessity of matching reproductive activity with environmental conditions conducive to spawning, fertilization, larval development, and recruitment. The implication is that \(N_e\) would be determined by the winners of this sweepstakes reproductive lottery, and if the number of winners was small, \(N_e\) would be orders of magnitude smaller than \(N\) (Hedgecock 1994; Hedgecock et al. 2007a).
The sweepstakes reproductive success hypothesis has been shown to be theoretically robust provided that there is high variance in reproductive success among families (Hedrick 2005). Knowledge of pedigree structure within groups of animals is important in a number of research areas in behavioral, ecological, and evolutionary genetics and in conservation biology. A cohort produced by a finite number of parents may include full and half sibling families, which can be detected by statistical methods. However, there are very few direct empirical studies of the large variance in reproductive success by studying the family structure within a year class. The sibling relationships among individuals coming from a single cohort were inferred by pairwise approaches in previous studies (Selkoe et al. 2006; Hedgecock et al. 2007b). However, valuable information may be lost in breaking the sampled individuals into pairs and considering each pair in isolation (Bulter et al. 2004; Wang 2004). Group approaches are expected to be more powerful than pairwise approaches because the former uses information of the multilocus genotypes of all sampled individuals to simultaneously assign them to sib groups (Blouin 2003; Bulter et al. 2004; Wang 2004).

Several reproductive features make striped bass likely to show patterns of sweepstakes reproductive success. *M. saxatilis* is an iteroparous, broadcast-spawning fish species with a type III survivorship curve. Characteristic mating behavior consists of a single female surrounded by several males, at or near the surface (Setzler et al. 1980). Eggs are broadcast loosely into flowing waters, and spawning by a single female is completed within a few hours per spawning season (Lewis & Bonner 1966; Salek et al. 2001). The pelagic eggs drift downstream, propelled by river currents, and hatch at a considerable distance from the spawning site (Bulak et al. 1997). The fecundity of *M. saxatilis* is highly correlated with weight, length, and age such that a large female could lay one million eggs in a single spawning event each spawning season (Lewis & Bonner 1966). Mortality rates in early life history stages of *M. saxatilis* are very high (Secor & Houde 1995; Bulak et al. 1997). Also, in the Santee-Cooper system nearly all of the early striped bass juveniles are found in a defined area that is accessible for sampling.
The first documented, self-sustaining, freshwater population of *M. saxatilis* is found in the Santee-Cooper system, South Carolina, USA (Scruggs 1957), which was landlocked because of the construction of two dams that impounded Lake Moultrie and Lake Marion. A fish ladder and a lock system have reconnected the lake populations with the Santee and Cooper River populations, but tagging studies show that this and all other South Carolina striped bass populations naturally avoid salt water (Bulak *et al.* 2004). In a subsequent study, Bulak *et al.* (1997) showed that mortality among striped bass egg cohorts was highly variable within a spawning season and that substantial recruitment of *M. saxatilis* came from relatively few river-spawn egg cohorts that were transported at the right time to high-quality nursery habitat in Lake Marion. In addition, levels of natural recruitment of *M. saxatilis* have been shown to fluctuate considerably among year-classes in the Santee-Cooper system (Bulak *et al.* 1997). All of these data indicate that the Santee-Cooper population is an attractive system for testing the hypothesis of sweepstakes reproductive success.

In the present study, yearly cohorts of juvenile striped bass collected from 1992 through 1994 from the Lake Marion nursery area (Diaz *et al.* 2000) were analyzed at seven polymorphic microsatellite loci. We hypothesized that the observed sweepstakes reproductive success would result in a low $N_e/N$ ratio as shown in the preliminary study by Diaz *et al.* (2000). Furthermore, since few adults were thought to contribute to each year class, we expected to find large full sibling families among the cohorts of juveniles. If the number of these large families was relatively small, then the temporal genetic variance among cohorts of juveniles would be significant.

**Materials and methods**

**Sample collections**

Striped bass juveniles ($N = 181$ to 225) were sampled with a $10.7 \times 1.8$ m beach seine with a $4.8$ mm mesh in 1992-1994 from the primary nursery ground near Santee State Park as described by Diaz *et al.* (2000). Because oxytetracycline-marked hatchery fish were stocked at the sampling site in
1992 and 1993, otoliths were checked for an oxytetracycline mark under a fluorescence compound microscope, and fish showing the mark were excluded from further analyses. In 1994, no hatchery fish were stocked so this procedure was unnecessary.

**DNA extraction and microsatellite genotyping**

Genomic DNA was isolated from fin clips or muscle tissue using a DNAzol Genomic DNA Isolation Reagent (DN 127, Molecular Research Center, Cincinnati, Ohio) after a proteinase K digest. Genotypes were determined at seven dinucleotide repeat microsatellite loci: SB108 (Skalski et al. 2006), SB91 (Roy et al. 2000), Labrax-6 (García de León et al. 1995), MSM1095, MSM1115, MSM1144, and MSM1168 (Couch et al. 2006). One of the two primers of each locus was 5’-end labeled with a fluorescent dye (HEX, FAM, or NED). These seven loci were amplified in three two-locus multiplexes (multiplex 1, MSM1095, and MSM1168, annealing temperature 58°C; multiplex 2, MSM1144, and Labrax-6, annealing temperature 51°C; multiplex3, MSM1115, and SB108, annealing temperature 56°C) and one single PCR (SB91, annealing temperature 58°C). Amplifications were performed on a Mastercycler ep Gradient S (Eppendorf) in a 12.5 µL reaction containing about 50 ng genomic DNA, 1.25 µL 10× buffer (New England Biolabs, Beverly, MA), 1.5mM MgCl₂, 0.2 mM of each dNTP, 0.04~0.24 µM each primer, and 0.5 U Taq DNA polymerase (New England Biolabs, Beverly, MA). Thermal cycling parameters for all amplifications were: 95 °C for 3 min, then 35 cycles each at 95 °C for 20 s, 20 s at the appropriate annealing temperature, and 72 °C for 30s, followed by 1 cycle of final elongation at 72 °C for 10 min. Amplified products were run on an ABI PRISM 3130 DNA analyzer with a CXR size standard (Promega, Madison, WI). Allele scoring was performed using GENEMAPPER software version 4.0 (Applied Biosystems, Foster City, CA).

**Data analysis**
For each annual cohort, observed heterozygosity and expected heterozygosity (Nei 1987) were calculated using Arlequin version 2.000 (Schneider et al. 2000). Number of alleles, and allelic richness were calculated by using FSTAT Version 2.9.3.2 (Goudet 1995). Allelic richness is a measure of the number of alleles that is relatively independent of sample size (Leberg 2002). Deviations from Hardy-Weinberg equilibrium were examined using an exact test based on a Markov Chain method (Guo & Thompson 1992) for each locus-cohort combination using GENEPOP (Raymond & Rousset 1995). The presence of genotypic disequilibrium between all pairs of loci in each cohort was tested using the log-likelihood ratio $G$ statistic implemented in FSTAT.

The fixation index $F_{ST}$ values (Weir & Cockerham 1984) based on the overall data set were estimated among all pairs of cohorts using Arlequin (Schneider et al. 2000). The statistical significance of $F_{ST}$ estimates was tested by 10,000 random permutations of individuals across populations. In addition, we conducted a hierarchical analysis of molecular variance (AMOVA; Excoffier et al. 1992) to test the significance of the portion of the total genetic variance imputable to cohorts using Arlequin. We also analyzed the genotypic differentiation of cohort samples with global and pairwise Fisher’s exact tests implemented with default Markov chain parameters in GENEPOP (Goudet et al. 1996).

An estimate of $N_e$ was obtained using the temporal method for overlapping generations as described in Jorde and Ryman (1995; 1996). The model requires a basic life table with information on age-specific survival rate ($l_i$) and birth rates ($b_i$). Age-specific survivorship, $l_i$, for the striped bass population in the Santee-Cooper system was estimated by Bulak et al. (1995). Average reproductive contribution ($b_i$) was estimated as fecundity at age $i$ (Bulak et al. 1995), and was standardized to result in a net reproductive rate $R_0 = \Sigma l_i b_i = 1$. Briefly, the $N_e$ values were calculated using the equation: $N_e = C / 2G\overline{F'}$. Equations to compute, $\overline{F'}$, the grand mean of temporal allele frequency change (sampling error corrected) among all adjacent cohorts and over all loci are found in Jorde & Ryman (1996). The correction factor ($C$) for overlapping generations was calculated by using 50 iterations of Equation 5 in Jorde & Ryman (1996). The mean generation length in years, $G$, was calculated using equation 10 in Jorde & Ryman...
(1996). In addition, the 95% confident interval (CI) for $\overline{F'}$ and $N_e$ was calculated as in Waples (1989) assuming a chi-square distribution. The estimated adult census size ($N = 362,000$) corresponded to the total number of adult fish estimated by Bulak et al. (1995).

The effective number of breeders ($N_b$) for each cohort was estimated by using the sibship assignment (SA) method as implemented in the software package Colony2 (Wang 2004). The SA method infers $N_b$ from the sibship frequencies estimated from a sibship assignment analysis using multilocus genotypes of a sample of offspring taken at random from a single cohort (Wang 2009).

The kinship analyses were carried out with a pairwise likelihood method implemented in the computer program KINGROUP (Konovalov et al. 2004) and with a pairwise score method implemented in the online software PEDIGREE 2.2 at http://herbinger.biology.dal.ca:5080/Pedigree. Both programs gave similar results. However, a significant number of false positive full sibs were included among the families identified by KINGROUP program. Therefore, only the results obtained using PEDIGREE were used in subsequent analyses. The pairwise score method of PEDIGREE uses a Markov chain Monte Carlo (MCMC) algorithm to move through the space of possible partitions, and locates a partition that maximizes an overall score based on the pairwise likelihood ratio of being full-sibs or unrelated (Smith et al. 2001; Butler et al. 2004). The presence of large full-sib families that are split in the partition with lower weight was detected by comparing the best partitions generated with weight 1, 2, and 5. The best full-sib partition generated with a weight of 2 was used in the analyses. To verify the accuracy of the family assignments, the 1992 and 1993 cohorts were combined, and the family analysis was repeated. No members of the 1993 cohort were assigned to the large families in the 1992 cohort and vice versa, demonstrating the accuracy of the family assignments. Potential false positives in large full-sib families were detected by comparing the list of observed offspring genotypes to the expected genotypes inferred from reconstructed parental genotypes.
Throughout the analysis, significance levels were adjusted for multiple simultaneous comparisons according to the sequential Bonferroni procedure using a global significance level of 0.05 (Rice 1989).

**Results**

In total, 603 recruits representing 3 yearly cohorts were analyzed at seven microsatellite loci with no missing data. All seven microsatellite loci were polymorphic in all cohorts. $H_o$ and $H_e$ ranged from 0.420 to 0.878 and from 0.419 to 0.860, respectively. Allelic richness per locus and cohort ranged from 4.00 at MSM1115 to 12.92 at SB06 (Table 1).

Only four locus-cohort combinations out of 21 showed significant deviation from Hardy-Weinberg equilibrium, with 3 occurring in cohort 1992 and 1 in cohort 1994 (Table 1). Linkage disequilibrium among pairs of loci was tested for 63 combinations in the three cohort samples with 12 pairs of loci, all in the 1992 cohort, exhibiting significant linkage disequilibrium after correction for multiple tests (Table 2). The results with the 1993 and 1994 cohorts indicated that no significant genetic linkage exists among the seven loci. Therefore, the linkage disequilibrium observed among the 1992 cohort suggests that only a small number of parents contributed to that cohort so that a limited number of combinations of unlinked alleles generated the apparent linkage disequilibrium. The observed deviation from Hardy-Weinberg equilibrium could also result from a low number of parents contributing to the 1992 and 1994 cohorts.

AMOVA revealed that a significant proportion of the observed genetic variation was attributable to differences among cohorts ($F_{ST} = 0.00605, P = 0.000$). Similar results were obtained with a global exact test of genotypic differentiation ($P = 0.000$). In addition, all three pairwise $F_{ST}$ values among cohorts were significant, as were the pairwise exact tests of genotypic differentiation (Table 3). These significant genetic differences among cohorts provide additional support for the idea that relatively small numbers of parents contributed to each of the year classes.
Differences in allele frequencies among annual cohorts can be used to estimate the effective population size (Jorde & Ryman 1995). However, in the case of striped bass, correction factors are needed to account for the presence of overlapping generations. The estimated correction factors ($C$) for overlapping generations and mean generation length ($G$) were 7.81 and 5.26, respectively (Bulak et al. 1995). The estimate of $\bar{F}$ was 0.0080 with a 95% CI of 0.0057 – 0.0120. Based on the equation: $N_e = \frac{C}{2G\bar{F}}$, an overall point estimate of $N_e$ was 93, with lower and upper 95% CI of 62 and 130, respectively. Dividing these values of $N_e$ by the estimated adult census size ($N = 362,000$) generated an estimate of $N_e/N = 2.56 \times 10^{-4}$ (95% CI: $1.71 \times 10^{-4} – 3.58 \times 10^{-4}$). Thus, the estimated $N_e$ values were roughly 4 orders of magnitude lower than the estimated adult census size.

The low $N_e$ values can be corroborated by calculating the effective number of breeders ($N_b$) contributing to each year class based on the distribution of full sibs within each cohort (Wang 2009). The estimate of $N_b$ for the 1992 cohort was 24 (95% CI: 14 – 42). Estimates of $N_b$ for the 1993 and 1994 cohorts were 40 (95% CI: 27 – 63) and 44 (95% CI: 29 – 68), respectively. Thus, even in 1993, which was a good recruitment year (personal communication, JS Bulak, South Carolina Department of Natural Resources, Eastover, SC), very few parents made significant contributions to the year class. Furthermore, all three estimates are consistent with a value of 93 for $N_e$ and with the estimates of $N_b$ obtained by Diaz et al. (2000) from changes in the allele frequencies of three nuclear DNA loci.

The best full-sib partition was generated for each cohort. When estimating the significance of groups through randomization trials, the smaller full-sib families of size 3 or below can rarely be distinguished from artefactual groupings of unrelated individuals (Pedigree Help Manual). Therefore, more emphasis was put on the larger full-sib families with at least four members in the analyses.

In the 1992 cohort, a total of 46 full-sib families were detected with the two largest full-sib families consisting of 27 and 19 individuals (Table 4, Figure 1). The 14 full-sib families that consisted of at least four members included 109 individuals out of a total of 181 fish tested (60%). In the 1993
cohort, 13 full-sib families with at least four members were detected with 7 individuals in the biggest full-sib family. Together, these families included 62 individuals out of a total of 197 individuals (31%). In the 1994 cohort, 22 full-sib families with at least 4 members were detected, and the biggest full-sib family comprised 6 individuals. Together, these families included 95 individuals out of a total of 225 individuals (42%). Thus, each year a small number of full sib families made up 31% to 60% of the year class (Table 4).

Discussion

The sweepstakes-mismatch process hypothesis was initially proposed to explain the low $N_e/N$ ratio observed in pelagic-spawning marine fishes and shellfishes. Although this hypothesis presents an attractive explanation for the lower than expected genetic diversity observed in marine populations, there is little empirical evidence that individuals experience the large variances in reproductive success proposed by the sweepstakes hypothesis (Flowers et al. 2002). The data presented here provide direct evidence that individual striped bass in the Santee-Cooper river system do experience large variance in reproductive success.

The observed reproductive variance was largest in the small 1992 cohort, with two full-sib families contributing 25% of the year class. However, even in 1993, which was a good recruitment year, 13 full-sib families were responsible for nearly one third of the year class (Table 4). These results are consistent with the recruitment dynamics of striped bass in the Santee-Cooper system. Bulak et al. (1997) showed that substantial recruitment in 1988 through 1990 came from relatively few eggs transported at the right time to high quality nursery habitat in Lake Marion. In some cases, periods of high relative survival and recruit production had their origins in weeks with relatively low levels of spawning (Bulak et al. 1997). These results suggested that only small numbers of striped bass spawning events are successful and that most spawning events did not make a significant contribution to recruitment.
Tagging studies indicate that the Santee-Cooper population of striped bass is completely isolated from other coastal populations (Bulak et al. 2004). Therefore, the differences that we observed in gene frequencies among samples cannot be explained by differences in migration from populations in nearby river systems. As a consequence, the temporal variation in allele frequencies that we observed is best explained by variance in reproductive success. Similar results have been obtained with several marine organisms (Hedgecock 1994; Ruzzante et al. 1996; Li & Hedgecock 1998; Moberg & Burton 2000; Planes & Lenfant 2002; Pujolar et al. 2006; Selkoe et al. 2006; Hedgecock et al. 2007b).

The severity of sweepstakes reproductive success varied among the three cohorts. The variance in full-sib family size was much larger in 1992 than in 1993 or 1994, and the proportion of individuals in full-sib families consisting of at least four members was also much higher in the 1992 cohort. Furthermore, the estimate of $N_b$ was lowest in the 1992 cohort, with the equivalent of only 12 breeding pairs creating the illusion of linkage disequilibrium. These results are consistent with the recruitment levels of these three cohorts. The annual recruitment of striped bass in the Santee-Cooper system varied by nearly two orders of magnitude from 1992 to 1994 (personal communication, JS Bulak, South Carolina Department of Natural Resources, Eastover, SC). Based on the catch of age II striped bass per 100,000 square feet of gill net, the recruitment index of cohort 1992 was only 2, indicating that 1992 was a poor recruitment year. However, the recruitment index was 179 and 41 for the 1993 and 1994 cohorts, respectively, suggesting that 1993 was an excellent recruitment year and that 1994 was a fair one. Variation in the recruitment levels of striped bass is caused chiefly by environmental conditions encountered by eggs and larvae. For example, water temperature, stream flow, and zooplankton prey availability all have been demonstrated to affect survival of striped bass eggs and larvae (Ulanowicz & Polgar 1980; Uphoff 1989; Secor & Houde 1995; Bulak et al. 1997). Our study suggests that the severity of the sweepstakes reproductive success varied depending on environmentally imposed mortality, which can be quite dynamic both within and among spawning seasons.
In addition to differences in the recruitment index, significant differences in microsatellite allele frequencies were observed among annual cohorts of striped bass, suggesting that different adults contribute to different cohorts. Evidence for this assertion comes from the fact that no members of the 1993 cohort were assigned to the large families from the 1992 cohort when the two sets of data were combined and analyzed together. In addition, the parental genotypes deduced for the large families were different for each year class. Thus, it does not appear that differences in reproductive success are due to differences in the reproductive fitness of a small number of adults.

The Santee-Cooper striped bass population is a managed population where the natural reproduction is supplemented with stocked fish from a hatchery to compensate for losses of reproductive potential due to fishing mortality. Each year, wild fish are captured and bred in the hatchery. A few weeks later, large numbers of the resulting progeny are returned to the river system. Thus, the stocking program with its inherently low numbers of contributing parents is mimicking the natural reproduction which has similar numbers of contributing parents. Since historic fishing pressure has resulted in 60 to 70% adult mortality (Bulak et al. 1995), the net effect of the hatchery program is probably similar to the natural reproduction that would have occurred in the absence of significant fishing pressure.

The low values of \(N_e\) and the small \(N_e/N\) ratio observed in striped bass are consistent with the low levels of genetic diversity observed in striped bass populations. Compared to other fish species, striped bass has extremely low levels of genetic polymorphisms in allozyme (Otto 1975; Rogier et al. 1985), mitochondrial DNA (Chapman 1990; Wirgin et al. 1993), and nuclear DNA (Leclerc et al. 1996; Diaz et al. 1998; Diaz et al. 2000). Even though striped bass populations are large, an extremely low effective population size would prevent the accumulation of genetic diversity. An exception would be the microsatellite DNA polymorphisms analyzed in the present study. Since mutation rates for microsatellite loci generally range from \(10^{-2}\) to \(10^{-6}\) (Lai & Sun 2003), new microsatellite alleles of the more rapidly mutating microsatellite loci would be present among the large families in each annual cohort. Thus, allelic variation at rapidly mutating microsatellite loci would accumulate in the population. In contrast,
mutations in microsatellite loci with slower mutation rates would rarely occur and would be unlikely to accumulate in the population. The data of Han et al. (2000) are consistent with this explanation since they found that only half of the microsatellite loci examined had observable allelic variation.

The use of a family analysis in this study was extremely powerful. Since multiple families of at least four members were observed even in the strong 1993 year class, it is clear that very few pairs of striped bass contribute to any particular year class, and as a result, $N_b$ estimations only ranged from 24 to 44 despite much larger variation in year class size. Furthermore, in-depth inspection of each large family allowed us to reconstruct the putative genotypes of the two parents who produced the family and identify potential false positive progeny. The presence of large two-parent families clearly indicates that striped bass engage in pair mating in the wild even though multiple males can be observed in the vicinity of spawning females. However, it is possible that these accessory males may fertilize a small number of the eggs produced by a spawning female.

Evidence of full and half siblings in cohorts of recruits has been obtained recently for marine organisms (Veliz et al. 2006; Selkoe et al. 2006; Hedgecock et al. 2007b). However, these studies addressed variance in reproductive success indirectly by inferring an excess of sibling relationships using pairwise approaches. The results presented here with striped bass suggest that group approaches are more powerful than pairwise approaches since we were able to accurately identify individual full-sib families and provide direct evidence for sweepstakes reproductive success. However, one limitation of the family analysis is that the sample size must be significantly larger than the actual value of $N_b$ so that significant family sizes can be obtained. Thus, the well defined sub-populations of striped bass and the small $N_b$ combine to make striped bass an excellent system for family analyses that may be difficult to replicate in other systems.

References


Hoarau G, Boon E, Jongma DN et al. (2005) Low effective population size and evidence for inbreeding in an overexploited flatfish, plaice 


Jorde PE, Ryman N (1996) Demographic genetics of brown trout (Salmo trutta) and estimation of effective population size from temporal change of allele frequencies. Genetics, 143, 1369-1381.


Wirgin II, Ong TL, Maceda L et al. (1993) Mitochondrial DNA variation in striped bass (Morone saxatilis) from Canadian rivers. Canadian Journal of Fisheries and Aquatic Sciences, 50, 80-87.

Wright S (1931) Evolution in Mendelian populations. Genetics, 16, 97-159.

Acknowledgements

We are grateful to Jeremy Dietrick for his technical assistance. We thank Dr. Jinliang Wang for assistance in running COLONY, Dr. Christophe Herbinger for assistance in running PEDIGREE. Thanks to Dr. James Bulak for sharing unpublished data. We also thank the two
anonymous reviewers for their comments. This work was supported in part by FISHTEC grant NOAA-NMFS RT/F-1 and NIH grant R25RR01854201.

**Author Information Box**

Jin-Xian Liu is a postdoctoral researcher with broad interest in ecology and evolutionary biology, especially of fishes. Dr. Bert Ely’s research group uses DNA technology to study the genetics of fish populations and human populations.

**Figure Legend**

Figure 1. Size distributions of full-sib families in each annual cohort.
Table 1 Summary statistics for microsatellite loci in three cohorts. Sample size (N), number of alleles (n), observed heterozygosity ($H_o$), expected heterozygosity ($H_e$), allelic richness ($A_R$) (based on minimum number of sample size of 181), inbreeding coefficients ($F_{IS}$) and associated $P$-value are presented.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cohort 1992, $N = 181$</th>
<th>Cohort 1993, $N = 197$</th>
<th>Cohort 1994, $N = 225$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$H_o$</td>
<td>$H_e$</td>
</tr>
<tr>
<td>MSM1095</td>
<td>7</td>
<td>0.702</td>
<td>0.714</td>
</tr>
<tr>
<td>MSM1168</td>
<td>5</td>
<td>0.619</td>
<td>0.576</td>
</tr>
<tr>
<td>MSM1144</td>
<td>11</td>
<td>0.801</td>
<td>0.798</td>
</tr>
<tr>
<td>Labrax-6</td>
<td>12</td>
<td>0.878</td>
<td>0.860</td>
</tr>
<tr>
<td>MSM1115</td>
<td>4</td>
<td>0.420</td>
<td>0.419</td>
</tr>
<tr>
<td>SB108</td>
<td>9</td>
<td>0.807</td>
<td>0.805</td>
</tr>
<tr>
<td>SB91</td>
<td>6</td>
<td>0.790</td>
<td>0.749</td>
</tr>
</tbody>
</table>

* significant departures from Hardy-Weinberg equilibrium (after sequential Bonferroni correction for multiple tests).
Table 2. *P* values for genotypic disequilibrium between all pairs of loci in each cohort.

<table>
<thead>
<tr>
<th>Locus Pair</th>
<th>Cohort 1992</th>
<th>Cohort 1993</th>
<th>Cohort 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSM1095 × MSM1168</td>
<td>0.00079*</td>
<td>0.03254</td>
<td>0.46429</td>
</tr>
<tr>
<td>MSM1095 × MSM1144</td>
<td>0.00079*</td>
<td>0.53571</td>
<td>0.86429</td>
</tr>
<tr>
<td>MSM1095 × Labrax-6</td>
<td>0.00159</td>
<td>0.22063</td>
<td>0.76429</td>
</tr>
<tr>
<td>MSM1095 × MSM1115</td>
<td>0.00079*</td>
<td>0.03016</td>
<td>0.97222</td>
</tr>
<tr>
<td>MSM1095 × SB108</td>
<td>0.00079*</td>
<td>0.05079</td>
<td>0.81587</td>
</tr>
<tr>
<td>MSM1095 × SB91</td>
<td>0.00397</td>
<td>0.85952</td>
<td>0.38889</td>
</tr>
<tr>
<td>MSM1168 × MSM1144</td>
<td>0.01984</td>
<td>0.66667</td>
<td>0.76032</td>
</tr>
<tr>
<td>MSM1168 × Labrax-6</td>
<td>0.00079*</td>
<td>0.00317</td>
<td>0.76905</td>
</tr>
<tr>
<td>MSM1168 × MSM1115</td>
<td>0.20873</td>
<td>0.02381</td>
<td>0.67460</td>
</tr>
<tr>
<td>MSM1168 × SB108</td>
<td>0.04048</td>
<td>0.39365</td>
<td>0.97143</td>
</tr>
<tr>
<td>MSM1168 × SB91</td>
<td>0.00238</td>
<td>0.30476</td>
<td>0.56667</td>
</tr>
<tr>
<td>MSM1144 × Labrax-6</td>
<td>0.00079*</td>
<td>0.47778</td>
<td>0.77063</td>
</tr>
<tr>
<td>MSM1144 × M1115</td>
<td>0.00159</td>
<td>0.72143</td>
<td>0.54048</td>
</tr>
<tr>
<td>MSM1144 × SB108</td>
<td>0.00159</td>
<td>0.24048</td>
<td>0.10714</td>
</tr>
<tr>
<td>MSM1144 × SB91</td>
<td>0.00079*</td>
<td>0.47698</td>
<td>0.00952</td>
</tr>
<tr>
<td>Labrax-6 × MSM1115</td>
<td>0.00079*</td>
<td>0.63413</td>
<td>0.90000</td>
</tr>
<tr>
<td>Labrax-6 × SB108</td>
<td>0.00079*</td>
<td>0.46825</td>
<td>0.17778</td>
</tr>
<tr>
<td>Labrax-6 × SB91</td>
<td>0.00079*</td>
<td>0.42063</td>
<td>0.08968</td>
</tr>
<tr>
<td>MSM1115 × SB108</td>
<td>0.00794</td>
<td>0.99286</td>
<td>0.50317</td>
</tr>
<tr>
<td>MSM1115 × SB91</td>
<td>0.00079*</td>
<td>0.95714</td>
<td>0.74841</td>
</tr>
<tr>
<td>SB108 × SB91</td>
<td>0.00079*</td>
<td>0.89444</td>
<td>0.25952</td>
</tr>
</tbody>
</table>

No. significant 12 0 0

* significant genotypic disequilibrium (after sequential Bonferroni correction for multiple tests).
Table 3. Pairwise $F_{ST}$/corresponding $P$-values (below diagonal) and Chi-square/corresponding $P$-values of pairwise Fisher’s exact tests (above diagonal)

<table>
<thead>
<tr>
<th></th>
<th>Cohort 1992</th>
<th>Cohort 1993</th>
<th>Cohort 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort 1992</td>
<td></td>
<td>infinity/0.0000</td>
<td>infinity/0.0000</td>
</tr>
<tr>
<td>Cohort 1993</td>
<td>0.0108/0.0000</td>
<td></td>
<td>36.540/0.0009</td>
</tr>
<tr>
<td>Cohort 1994</td>
<td>0.0068/0.0000</td>
<td>0.0015/0.02762</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Results of kinship analysis.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>N*</th>
<th>RI</th>
<th>n$_f$</th>
<th>n$_{≥4}$</th>
<th>N$_{≥4}$</th>
<th>N$_{≥4}$/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>181</td>
<td>2</td>
<td>46</td>
<td>14</td>
<td>109</td>
<td>0.60</td>
</tr>
<tr>
<td>1993</td>
<td>197</td>
<td>179</td>
<td>71</td>
<td>13</td>
<td>62</td>
<td>0.31</td>
</tr>
<tr>
<td>1994</td>
<td>225</td>
<td>41</td>
<td>74</td>
<td>22</td>
<td>95</td>
<td>0.42</td>
</tr>
</tbody>
</table>

*Symbols: N, number tested; RI, Recruitment Index (catch of age II striped bass per 100,000 square feet of gill net); n$_f$, number of full-sib families; n$_{≥4}$, number of full-sib families of size ≥ 4; N$_{≥4}$, number of individuals in full-sib families of size ≥ 4; N$_{≥4}$/N, proportion of individuals in full-sib families of size ≥ 4.