Molecular Ecology (2004) **13**, 2947–2958 doi: 10.1111/j.1365-294X.2004.02303.x

# Conservation genetics and demographic history of the endangered Cape Fear shiner (*Notropis mekistocholas*)

E. SAILLANT, J. C. PATTON, K. E. ROSS and J. R. GOLD
*Center for Biosystematics and Biodiversity, Texas A&M University, College Station, Texas, 77843–2258, USA*

## Abstract

**We examined allelic variation at 22 nuclear-encoded markers (21 microsatellites and one anonymous locus) and mitochondrial (mt)DNA in two geographical samples of the endangered cyprinid fish *Notropis mekistocholas* (Cape Fear shiner). Genetic diversity was relatively high in comparison to other endangered vertebrates, and there was no evidence of small population effects despite the low abundance reported for the species. Significant heterogeneity (following Bonferroni correction) in allele distribution at three microsatellites and in haplotype distribution in mtDNA was detected between the two localities. This heterogeneity may be due to reduced gene flow caused by a dam built in the early 1900s. Bayesian coalescent analysis of microsatellite variation indicated that effective population size of Cape Fear shiners has declined in recent times (11–25 435 years ago, with highest posterior probabilities between 126 and 2007 years ago) by one–two orders of magnitude, consistent with the observed decline in abundance of the species. A decline in effective size was *not* indicated by analysis of mtDNA, where sequence polymorphism appeared to carry the signature of an older expansion phase that dated to the Pleistocene (~12 700 > 1 million years ago). Cape Fear shiners thus appear to have undergone an expansion phase following a glacial cycle but to have declined significantly in more recent times. These results suggest that rapidly evolving markers such as microsatellites may constitute a suitable tool when inferring recent demographic dynamics of populations.**

*Keywords*: Cape Fear shiners, conservation genetics, demographic history, effective population size, microsatellites, mtDNA

*Received 19 March 2004; revision received 7 June 2004; accepted 2 July 2004*

## Introduction

Efforts to manage and conserve threatened and endangered populations of freshwater fishes have been supported in many ways by deployment of molecular markers to resolve taxonomic problems and describe patterns of genetic variations both within and among lineages (Vrijenhoek 1998) and to infer evolutionary and demographic history of species under decline (Moran 2002). The latter has been improved greatly by the recent advent of genealogical modelling of the coalescent process, permitting comparisons of statistics calculated from a data sample with the distribution of the statistics in simulated samples (Beaumont 1999). Coalescent-based analyses thus allow, in theory, robust inference even when limited sample sizes are available, as it is often the case for endangered species.

In this study, we utilized a suite of molecular markers to describe patterns of genetic variation and infer demographic history of the Cape Fear shiner, *Notropis mekistocholas* (Snelson 1971), a small, herbivorous cyprinid fish (minnow) endemic to the Cape Fear river basin in the east-central Piedmont region of North Carolina (USA) and one of the 187 species of freshwater and diadromous fishes in the southern USA recognized as extinct, endangered or threatened (Warren *et al*. 2000). Cape Fear shiners are reported to occur in only five localities in the Cape Fear drainage (North Carolina Wildlife Resources Commission 1995), but in recent years have not been found in two of these localities, both in the Haw River (D. Rabon, US Fish and Wildlife Service, personal communication). In addition, very recent surveys at a third locality, the Deep River above the dam at High Falls in Moore County, have

Correspondence: Eric Saillant. Fax: (1) 979 8454096; E-mail: esaillant@tamu.edu

yielded only a single Cape Fear shiner (A. Howard, North Carolina State University, personal communication). While it may be that Cape Fear shiners have always existed in low numbers (Snelson 1971), Pottern & Huish (1986, 1987) reported declines in abundance and range for the species that were not evident for other sympatric taxa. Cape Fear shiners are currently designated as endangered (Federal Register 1987; Hilton-Taylor 2000; http://www.unep-wcmc.org/index.html? http://www.unep-wcmc.org/species/index.htm~main); potential causes for their decline include changes in stream morphology and flow (US Fish and Wildlife Service 1988) and reduced habitat quality and availability (Howard 2003).

To date, allelic variation at 23 genetic markers (21 microsatellites, one anonymous nuclear-encoded locus and mtDNA) has been examined among samples collected in 1998 from the three localities where Cape Fear shiners have been captured in the last decade (Burridge & Gold 2003; Gold *et al*. 2004). Although relatively high levels of genetic diversity and no evidence of small population effects or inbreeding were found, a recent reduction in effective population size was suggested by the ratio of the number of (microsatellite) alleles to the range in allele size (the *M*-test of Garza & Williamson 2001). However, sample sizes employed in these studies were small (10–15 individuals per locality), due in large part to the rarity and difficulty in capturing Cape Fear shiners. Here, we report results from (noninvasive) sampling of additional specimens at the two localities where Cape Fear shiners still occur in relative abundance. We assayed the same set of 23 genetic markers used by Gold *et al*. (2004). Our objectives were to (i) further examine genetic diversity and effective population size in extant Cape Fear shiners in order to ascertain whether small population effects were evident; (ii) assess more rigorously whether genetic heterogeneity existed between the sample localities; and (iii) employ recently developed, coalescent approaches (Kuhner *et al*. 1998; Beaumont 1999; Storz & Beaumont 2002) and mismatch distributions (Rogers & Harpending 1992) to infer the demographic dynamics of Cape Fear shiners at each locality.

## Materials and methods

### *Study sites and sample collection*

Adult Cape Fear shiners were sampled during the spring and summer of 2002 from two of the three localities sampled in 1998. Full details of the sampling localities may be found in Burridge & Gold (2003); a map of the area, including sampling localities, is shown in Fig. 1. The two localities sampled in 2002 were Confluence ($n = 29$) and High Falls ($n = 26$); Confluence is located 13.5 miles downstream from High Falls and the two localities are separated by the Carbonton Dam built in 1922 (D. Rabon,

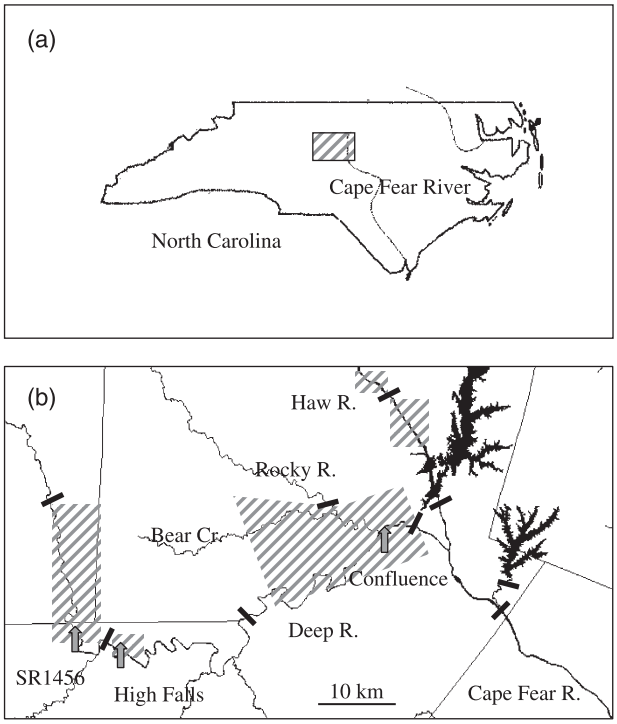


**Fig. 1** Sample localities for Cape Fear shiners (*Notropis mekistocholas*) in the Cape Fear River drainage (North Carolina, USA). Shaded areas represent the distributional limits of the species. R, River, C, Creek, black bar, dam.

US Fish and Wildlife Service, personal communication). Small clips from the lower lobe of the caudal fin were removed, stored in 95% ethanol and returned to College Station, where they were stored at −80 °C. All the fin-clipped fish swam away with no noticeable effect. We also included in this analysis the 15 samples from both Confluence and High Falls obtained in 1998; thus the overall sample sizes for Confluence and High Falls were 44 and 41, respectively.

### *Genetic assays*

All specimens were genotyped at 22 nuclear-encoded genetic markers. These markers included 21 microsatellites and one anonymous locus (*Nme*18A6.0) that coamplifies with microsatellite *Nme*18A6.158. Details of polymerase chain reaction (PCR) amplification, electrophoresis, and scoring of PCR products may be found in Burridge & Gold (2003) and Gold *et al*. (2004). Variation in mtDNA was assessed by directly sequencing fragments within the ND-5 [258 base pairs (bp)] and ND-6 (367 bp) subunits of mitochondrially encoded NADH dehydrogenase. PCR primer pairs, amplification conditions and sequence alignment procedures may be found in Gold *et al*. (2004). The partition test of Farris *et al*. (1995a,b), as implemented in PAUP 4.0b10 (Swofford 1998), was used to assess whether evolutionary rates of the two fragments of mtDNA were

homogeneous. The proportion of randomly generated partitions (100 replicates) that showed a greater incongruence index (Mickevich & Farris 1981) than the sequenced fragments was 0.52; consequently, the hypothesis of congruence was accepted and the sequences from both mtDNA fragments were combined into unique haplotypes for all subsequent analyses.

### *Data analysis*

Allele (nuclear) and haplotype (mtDNA) frequencies, number of alleles and haplotypes, allele richness (a measure of the number of alleles independent of sample size) and unbiased gene (microsatellites) diversity were computed using FSTAT version 2.9.3 (Goudet 1995). MtDNA number of polymorphic sites and nucleon and nucleotide diversity were computed in ARLEQUIN (Schneider *et al*. 2000). Gene and nucleon diversity were estimated after Nei (1987). Homogeneity of allele richness and of gene diversity between pairs of samples was tested using Wilcoxon signed-rank tests. Conformity of genotype proportions to Hardy–Weinberg equilibrium expectations within samples was measured for each nuclear-encoded locus as Weir & Cockerham 1984) *f*, calculated using FSTAT (Goudet 1995). Estimates for individual nuclear-encoded microsatellites were combined to generate a weighted estimate of *f* over all microsatellites, following recommendations in Weir & Cockerham (1984). Probability of significance of departure of genotype proportions from Hardy–Weinberg expectations was assessed by a Markov chain method (Guo & Thompson 1992), as implemented in GENEPOP version 1.2 and using 5000 dememorizations, 500 batches and 5000 iterations per batch (Raymond & Rousset 1995). Genotypic disequilibrium between pairs of microsatellites was assessed via an exact test, as implemented in GENEPOP and employing the same parameters as above for the Markov chain.

Homogeneity of allele distributions between temporal samples at each locality and between localities was measured as Weir & Cockerham 1984) θ, using FSTAT; probability of significance of the observed differences in allele distributions was assessed by an exact test, as implemented in GENEPOP (same parameters as above for the Markov chain). Homogeneity of mtDNA haplotype distributions (temporal and spatial, as above) was measured by $F_{ST}$ values based on genetic distances between sequences of mtDNA haplotypes and using Kimura (1980) two-parameter distance estimate. Significance of the distance estimates was tested via 10 000 permutations of haplotypes, as implemented in ARLEQUIN (Schneider *et al*. 2000). Sequential Bonferroni correction (Rice 1989) was applied when appropriate for all multiple tests performed simultaneously. A network of mtDNA haplotypes was constructed following Templeton *et al*. (1992), using the package TCS version 1.13.

### *Demographic history*

The demographic history of Cape Fear shiners at each locality was explored using microsatellite data and a Bayesian coalescent-based approach (Beaumont 1999; Storz & Beaumont 2002). A population changing in size exponentially was considered, and demographic parameters were inferred based on microsatellite genotypes and assuming a stepwise mutation model. The parameters estimated were $N_0$, $N_1$, μ and $t_a$. $N_0$ and $N_1$ are the effective number of chromosomes at sampling and at the beginning of an expansion/decline phase, respectively; μ is the average mutation rate over loci per generation; and $t_a$ is the time since the beginning of an expansion/decline phase. The ratio $r = N_0/N_1$ is expected to be $< 1$ in a declining population, equal to 1 in a stable population and $> 1$ in an expanding population (Beaumont 1999). Monte Carlo Markov chain (MCMC) simulations were used to estimate the posterior distribution of the genealogical history (mutational and coalescent events) and demographic parameters (growth rate and time since expansion) for each locality. Three chains ($10^9$ steps per chain, with the first $5 \times 10^8$ steps discarded to remove possible bias due to the starting values of the parameters) were run using the program MSVAR1.3.EXE, kindly provided by M. Beaumont of the University of Reading, UK. The three chains were based on different starting parameters in order to confirm convergence of the MCMC. The means of the prior distributions of means $N_0$, $N_1$ and $t_a$ were set (on a $\log_{10}$ scale) to 4 and their standard deviations (SD) were set to 3 in order to provide support for a wide a priori range for the parameters. The SDs of the prior distributions of the variance among microsatellites for these three parameters were set to 0.5 so that expected ratios for pairs of microsatellites would be of approximately fivefold under the prior (Storz & Beaumont 2002). Mean and SD for the prior distribution of mean μ were set to –3.3 and 0.5 so that values in the range $5 \times 10^{-3}$–$5 \times 10^{-5}$ would have reasonable support (Storz & Beaumont 2002); this prior thus utilized values that are widely assumed in demographic analyses (Heath *et al*. 2002; Storz & Beaumont 2002). The SD of the prior distribution of the variance of μ among microsatellites was set to 2 so that an expected ratio of mutation rates for pairs of microsatellites of about 700-fold would be frequent under the prior (Storz & Beaumont 2002). The generation time of Cape Fear shiners is not known; however, life-history data on other shiners (Harrell & Cloutman 1978; Matthews & Heins 1984; Cloutman & Harrell 1987) indicate that a generation time between 1 and 3 years would be appropriate. Consequently, time since expansion/decline was estimated considering generation time between 1 and 3 years. Estimates of current effective population size ($N_e$) and confidence limits were calculated by using the relationship $N_e = N_0/2$.

The mtDNA haplotype sequence data also were used to infer demographic history at each locality. Neutrality of the mtDNA sequences was tested by means of Fu's (1997) $F_S$ statistic, as implemented in ARLEQUIN, and Fu & Li's (1993) $D^*$ and $F^*$ statistics, as implemented in DNASP (Rozas *et al*. 2003). Significance of $F_S$ was assessed by 10 000 randomizations; significance of $D^*$ and $F^*$ was assessed using 10 000 coalescent simulations based on the observed number of segregating sites in each sample. Demographic history of mtDNA was first examined using the maximum-likelihood coalescent approach of Kuhner *et al*. (1998), as implemented in the program FLUCTUATE version 1.4, available at http://evolution.genetics.washington.edu/lamarc/fluctuate.html. FLUCTUATE utilizes a model where a single population has been expanding (or declining) exponentially and provides estimates of the parameters $\Theta$ and $g$. $\Theta$ is defined as $4N_1\mu$ (where $N_1$ is the 'current' effective population size and $\mu$ is the neutral mutation rate per site), and $g$ is the exponential growth rate of the population (in $1/\mu$ generations). Watterson (1975) segregating sites estimate was used as the initial estimate of $\Theta$ for each run. FLUCTUATE requires input of a transition/transversion ratio to be applied during estimation. To determine this ratio, the model of DNA substitution that best fitted the mtDNA sequence data was determined based on the likelihood-ratio test statistic, as implemented in MODELTEST version 3.06 (Posada & Crandall 1998). The best model [H–Kishino–Yano (HKY) model of Hasegawa *et al*. 1985] involved variable base frequencies and a transition/transversion ratio of 13 for the mtDNA sequences assayed. This ratio was used to parameterize the mutation model implemented by FLUCTUATE. One hundred short chains (1000 steps per chain with a sampling increment of 100) followed by 10 long chains (500 000 steps per chain with a sampling increment of 500) were employed to ensure convergence of the MCMC. Estimates of time (in generations) since a 'population' differed by one and two orders of magnitude from the 'current' effective size were determined by using the formula $t_s = (\ln(N_t/N_1))/(g\mu)$, where $N_1$ and $N_t$ are the effective population sizes at present ($N_1$) and $t$ generations ago ($N_t$) ago, respectively. Given the wide range of mutation rates that have been estimated for coding regions of mtDNA in fish (e.g. 1.2% for NADH 2 in several geminate species pairs of fishes according to Bermingham *et al*. 1997 to *c*. 5% for ND5/6 in salmonids according to Doiron *et al*. 2002), values of $\mu$ employed in the estimation were 1.0 and 5.0% $MY^{-1}$. Demographic history based on mtDNA sequence variation was then examined by mismatch-distributions analysis (Rogers & Harpending 1992). The sum-of-squared-difference statistic (SSD) was used to compare the observed (mismatch) distribution of pairwise differences between individuals to that expected under demographic expansion; significance of SSD was assessed by parametric bootstrap resampling (Schneider & Excoffier 1999), as implemented in ARLEQUIN (10 000 resamplings).

## Results

Detailed summary statistics, including allele and genotype distributions for the 22 nuclear-encoded loci and haplotype distributions for mtDNA, for each locality and each year sampled may be obtained from the corresponding author. Temporal variation, i.e. between sample years, in allele frequency over all 22 (21 microsatellites) nuclear-encoded loci was nonsignificant in both Confluence ($P = 0.176$) and High Falls ($P = 0.790$). Probability values for individual loci ranged from 0.011 (*Nme*30D2-120 in Confluence) to unity (*Nme*30D2-120 and *Nme*24B6-191 in High Falls); none of the probability values were significant following Bonferroni correction. MtDNA haplotype frequencies also did not differ significantly between sample years ($P = 0.134$, Confluence; $P = 0.835$, High Falls).

Summary data of variation for each nuclear-encoded locus at each sample locality, pooled across years, are presented in Appendix I. All nuclear-encoded loci were polymorphic. Number of alleles sampled per locus per locality ranged from two (*Nme*24B6.191, both localities; *Nme*30D2.120, High Falls) to 14 (*Nme*30F12.208, both localities); allelic richness per locus averaged 5.5 and ranged from 1.76 (*Nme*24B6.191, Confluence) to 8.4 (*Nme*2B10.232, Confluence). Gene diversity per locus averaged 0.703 and ranged from 0.128 (*Nme*24B6.191, Confluence) to 0.889 (*Nme*2D5.117, High Falls). No significant difference in allelic richness ($P = 0.81$) or gene diversity ($P = 0.62$) was detected in pairwise comparisons between localities. Tests of conformity to Hardy–Weinberg equilibrium expectations were nonsignificant following Bonferroni correction except for *Nme*12D8.125, where a highly significant deficit in heterozygotes was observed for both localities ($F_{IS}$ values were 0.372 and 0.672, respectively, for Confluence and High Falls). The heterozygote deficiency may reflect occurrence of null alleles; consequently, *Nme*12D8.125 was omitted from all subsequent analyses. A total of 22 of 462 (pairwise) tests of genotypic disequilibrium were significant ($P < 0.05$) prior to Bonferroni correction; of these, two (*Nme*2B10.166/*Nme*2B10.232, Confluence; and *Nme*27F8.230/*Nme*2B10.232, High Falls) remained significant after correction. A total of 19 distinct mtDNA haplotypes were detected (Fig. 2). Nucleon diversity values were $0.80 \pm 0.06$ (Confluence) and $0.85 \pm 0.03$ (High Falls), and nucleotide diversity values were $3.6 \times 10^{-3}$ (Confluence) and $4.3 \times 10^{-3}$ (High Falls).

### *Geographic structure*

Significant heterogeneity ($P < 0.05$) of allele frequencies between the two localities was detected at 14 of the

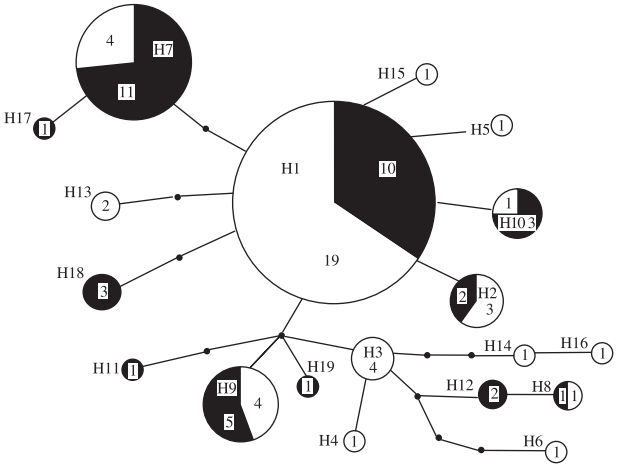


**Fig. 2** Network of haplotypes of Cape Fear shiners (*Notropis mekistocholas*) sampled within the Cape Fear river drainage. Haplotypes (H) found in Confluence are in the white area (overall sample size $n = 44$); those in the black area are for High Falls (overall sample size $n = 41$). ● missing haplotype. GenBank Accession nos for composite haplotypes are as follow (ND5–ND6): H1 (AY396545–AY396555), H2 (AY396546–AY396555), H3 (AY396547–AY396556), H4 (AY396548–AY396556), H5 (AY396545–AY396557), H6 (AY396549–AY396558), H7 (AY396550–AY396559), H8 (AY396551–AY396560), H9 (AY396552–AY396556), H10 (AY396545–AY396561), H11 (AY396553–AY396556), H12 (AY396554–AY396560), H13 (AY422787–AY422791), H14 (AY396547–AY422792), H15 (AY422788–AY396555), H16 (AY396545–AY422792), H17 (AY422789–AY396559), H18 (AY422790–AY396562), H19 (AY396545–AY422793).

nuclear-encoded loci prior to Bonferroni correction; three of these (*Nme*2B10.232, *Nme*5G4.175, and *Nme*6A7.93) remained significant after correction. The overall degree of divergence was relatively low ($F_{ST}$ of 0.014, $P < 0.0001$). Significant heterogeneity ($F_{ST} = 0.029$, $P = 0.035$) between localities in mtDNA haplotype frequencies also was detected. The mtDNA haplotype topology (Fig. 2) revealed no evident phylogeographical separation (reciprocal monophyly) of haplotypes within either locality.

### *Demographic history*

The anonymous, nuclear-encoded locus (*Nme*18A6.0) that coamplifies with *Nme*18A6.158 was removed from this analysis, as variation at this locus could not be assumed to follow the stepwise mutation model assumed in MSVAR computations. In addition, alleles that differed by only a single base pair were detected at five of the (20) remaining loci. Such alleles were very infrequent at three of the loci (one at *Nme*18A6.158 and *Nme*5B10.211, and three at *Nme*24B6.211). Genotypes of the corresponding individuals (total of four) at these three loci were removed from the data set and only genotypes of the remaining individuals at each locus were used in the analysis. A total of 19 and 84 such alleles were found at microsatellites *Nme*7B9.190 and *Nme*27F8.230, respectively. These two microsatellites were omitted from the analysis; the final data set thus included 18 microsatellites.

**Table 1** Summary statistics for posterior distributions of the parameters μ, $N_e$, $r$ and $t_a$. Estimates were based on variation at 18 nuclear-encoded microsatellites

| | Mode | 0.05 quartile | 0.95 quartile |
|---|---|---|---|
| Confluence | | | |
| μ | $2.5 \times 10^{-4}$ | $4.5 \times 10^{-5}$ | $1.7 \times 10^{-3}$ |
| $N_e$ | 513 | 54 | 4463 |
| $\text{Log}_{10}(r)$ | −0.97 | −0.72 | −1.89 |
| $t_a$ (year)† | 669–2007 | 8478–25 435 | 62–185 |
| High Falls | | | |
| μ | $2.5 \times 10^{-4}$ | $4.2 \times 10^{-5}$ | $1.6 \times 10^{-3}$ |
| $N_e$ | 85 | 7 | 1235 |
| $\text{Log}_{10}(r)$ | −2.08 | −1.20 | −2.87 |
| $t_a$ (year)† | 126–379 | 2108–6323 | 11–33 |

†Range covering generation times between 1 and 3 years.

Bayesian coalescent analysis revealed a negative posterior distribution of $\log_{10}(r)$ values, consistent with a historical decline in effective size among Cape Fear shiners at both localities (Table 1). The corresponding (estimated) reduction in effective size ranged from one to three orders of magnitude, with modes between 0.97 (Confluence) and 2.08 (High Falls) (Table 1). Estimates of $N_e$ (the current effective population size) ranged from 7 to 4463, with modes at 513 (Confluence) and 85 (High Falls). Corresponding estimated average mutation rates for the 18 microsatellites analysed were between $4.5 \times 10^{-5}$ and $1.7 \times 10^{-3}$ (mode = $2.5 \times 10^{-4}$). Given the assumed generation-time range of 1–3 years, the estimated time since decline for Confluence ranged from 62 to 25 435 years ago (mode between 669 and 2007 years ago, depending on the estimate of generation time); for High Falls the time since decline ranged from 11 to 6323 years ago (mode between 126 and 379 years ago).

Results of neutrality tests for mtDNA sequences are shown in Table 2. Fu's $F_S$ statistic indicated a significant departure from neutrality for the sample from Confluence ($P = 0.007$) but not for the sample from High Falls ($P = 0.200$). Fu & Li's $D^*$ and $F^*$ statistics were nonsignificant ($P > 0.05$) for both samples. According to Fu (1997), these results indicate that population expansion or 'hitchhiking', rather than background selection, accounts for the significant departure from neutrality detected in the *Confluence* sample by $F_S$. The hypothesis of population expansion under the assumption of selective neutrality was examined further by the coalescent approach of Kuhner *et al*. (1998) and by mismatch-distribution analysis (Rogers

**Table 2** Fu's (1997) $F_S$ and Fu & Li's (1993) $D^*$ and $F^*$ measures of selective neutrality; probabilities of significance were estimated from randomizations ($F_S$) or coalescent simulations ($D^*$ and $F^*$)

| | $F_S$ | $P_{(Fs < Fsobs)}$ | $F^*$ | $P_{(F^* < F^*obs)}$ | $D^*$ | $P_{(D^* < D^*obs)}$ |
|---|---|---|---|---|---|---|
| Confluence | –5.60 | 0.007 | –1.59 | 0.079 | –1.17 | 0.099 |
| High Falls | –1.99 | 0.200 | –0.73 | 0.244 | –0.47 | 0.379 |

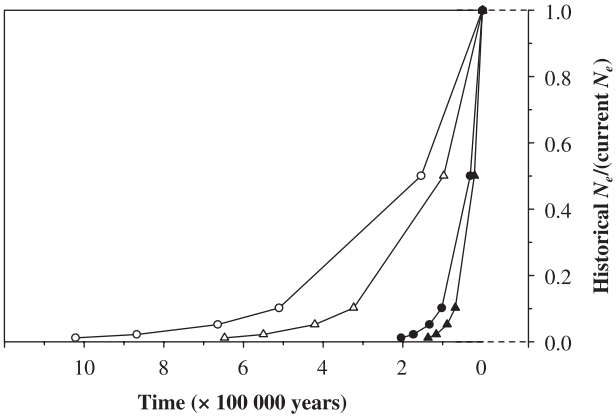


**Fig. 3** Plot of the ratio of historical to present-day effective population size, based on coalescent simulations carried out with mitochondrial DNA sequence data for Cape Fear shiners from two localities within the Cape Fear River drainage (Confluence, $n = 44$, triangles; High Falls, $n = 40$, circles). Per site mutation rates assumed were 1.0% (open triangles and circles) and 5% (filled in triangles and circles) per million years.

& Harpending 1992). FLUCTUATE runs indicated positive point estimates for the growth parameter ($g$) for both localities: $g$ (± standard deviation) = 672.5 ± 245.0 for Confluence and 450.4 ± 193.0 for *High Falls*. Estimates of Θ (± standard deviation) were 0.020 ± 0.006 (Confluence) and 0.011 ± 0.002 (High Falls). The growth parameter, although positive, corresponds to relatively slow growth where population size would have been 10% of 'current' size between 68 475 and 511 248 years ago (Fig. 3).

The mismatch distribution of mtDNA sequence differences among individuals at each locality (Fig. 4) did not differ from that expected under an expansion model; the sum-of-squared-difference (SSD) statistics were 0.020 ($P_{SSD}$ = 0.236) for Confluence and 0.018 ($P_{SDD}$ = 0.157) for High Falls. The time at which such demographic expansion would have occurred was estimated using the relationship $\tau = 2ut$, where $\tau$ is the crest of a unimodal mismatch distribution, $u$ is the mutation rate per generation of the DNA region under study and $t$ is the time in generations since demographic expansion (Rogers & Harpending 1992). Values of $\tau$ (Fig. 4) were obtained from ARLEQUIN; the parameter $u$ was estimated as $m_T\mu$, where $\mu$ is the mutation rate per nucleotide and $m_T$ is the number (625) of nucleotides assayed; $\mu$ values were set to 1.0 and 5.0% MY$^{-1}$ as above. Estimates of the time in years since demographic expansion ranged from 92 320 (Confluence, $\mu$ = 5.0% MY$^{-1}$) to 522 240 (High Falls, $\mu$ = 1% MY$^{-1}$). Considering the upper and lower bounds of $\tau$ (Fig. 4), the time since demographic expansion could vary between 12 704 and 1126 400 years ago.

## Discussion

The first objective of this study was to further assess genetic diversity in extant Cape Fear shiners in order to ascertain whether small population effects were evident. Small populations often exhibit reduced levels of genetic diversity and increased probability of identity by descent or inbreeding (O'Brien *et al*. 1996). Reduced genetic diversity may impact negatively a population's capability to respond to future environmental change or challenge, and increased inbreeding can expose deleterious recessives that can lead to reduction in overall fitness (Frankham 1995; Lynch *et al*. 1995; Frankham *et al*. 2002). For the two populations

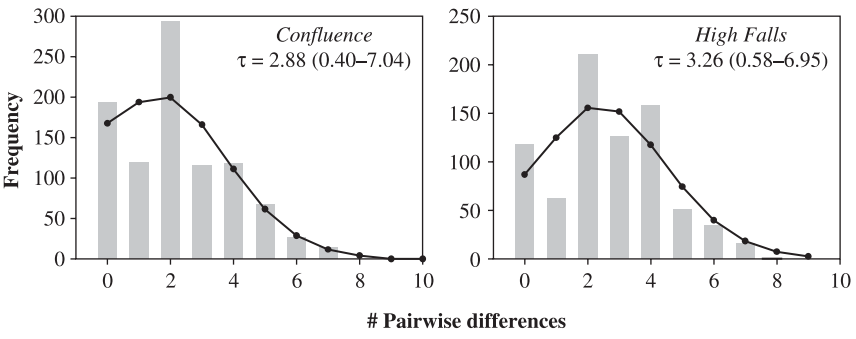


**Fig. 4** Mismatch distributions based on mitochondrial DNA sequences of Cape Fear shiners from two localities within the Cape Fear River drainage. Bars represent observed frequency of number of pairwise differences between sequences; lines represent the expected distribution assuming demographic expansion. Estimated values of $\tau$ (crest of the mismatch distribution) and its upper and lower bounds, generated by the bootstrap percentile method, are given for both samples.

examined here, the number of alleles over all nuclear-encoded loci averaged 8.18, while gene diversity ($H_E$) per locus averaged 0.701. These averages are nearly the same as or higher than those (7.5 alleles/locus, gene diversity of 0.46) reported by DeWoody & Avise (2000) for freshwater fishes, and considerably higher than averages for other critically endangered vertebrates, including the North Atlantic right whale (Waldick *et al*. 2002), Galápagos penguin (Akst *et al*. 2002), Gilbert's potoroo (Sinclair *et al*. 2002) and Delmarva fox squirrel (Lance *et al*. 2003). In these endangered species, the number of alleles per (microsatellite) locus ranged from 2.0 to 3.2, and gene diversities ranged from 0.038 to 0.457. Levels of mtDNA diversity in Cape Fear shiners also were relatively high. Nineteen mtDNA haplotypes were uncovered among the 85 individuals surveyed, and nucleon diversity (the probability that any two individuals drawn at random will differ in mtDNA haplotype) in both populations was greater than 0.80. Collectively, these results indicate that levels of genetic diversity in Cape Fear shiners are relatively high, despite current rarity of the species and its reported decline(s) in abundance and range (Pottern & Huish 1986, 1987). There was also no indication of significant inbreeding in the two populations, as estimated $F_{IS}$ values for the nuclear-encoded loci were less than 0.05 and none differed significantly from zero.

One caveat to the above is that historical levels of genetic diversity in Cape Fear shiners may have been significantly higher than those found here. The two Cape Fear shiner populations appear to have undergone significant effective-size reduction in the recent past (see below), and as pointed out by Hauser *et al*. (2002) changes in genetic (allelic) diversity may be in part a function of the history of population-size reduction and may not always be detectable, even though the population is in, or has experienced, a recent decline. Overall, however, results presented here suggest that high levels of genetic diversity are still present in current Cape Fear shiner populations.

The second objective of the study was to assess more rigorously whether genetic heterogeneity existed among Cape Fear shiners within the drainage. Burridge & Gold (2003) and Gold *et al*. (2004) reported significant allele-frequency differences at one or more nuclear-encoded loci between samples taken in 1998 at Confluence and High Falls. For most loci, however, the allele-frequency differences were not significant following Bonferroni adjustment. With the increased sample sizes obtained for this study, significant heterogeneity between the two localities was detected both before (14 microsatellites) and after (three microsatellites) Bonferroni adjustment for the nuclear-encoded loci and for mtDNA as well.

Given that Cape Fear shiners at the two localities differ genetically, it then becomes of interest to ask whether the two populations were separated historically or whether the dam constructed between the two localities in the early 1900s could have impeded gene flow sufficiently to allow genetic differences to accumulate. The former would suggest that the two populations were on independent evolutionary trajectories, whereas the latter might reflect the impact of human activities. Historical separation and the effective absence of gene flow between the two localities would be expected, in time, to lead to reciprocal monophyly of mtDNA haplotypes (Avise *et al*. 1987). The network of haplotypes of all Cape Fear shiner mtDNA haplotypes, however, revealed no evident phylogeographical separation (reciprocal monophyly) of haplotypes within either locality. We then asked if genetic drift alone could account for genetic differences observed between the two localities given the time period involved (1922 to present). We used the equation (from Hedrick 1999)

$$F = 1 - \left(1 - \frac{1}{2N}\right)^t$$

where $F$ is $F_{ST}$ among idealized populations under an infinite-island model (and assuming no migration), $N_e$ is effective size and $t$ is time in generations. For $F_{ST}$, we used 0.014 (the weighted $F_{ST}$ over all loci examined) and for $N_e$ we used the harmonic means of the lower ($N_e$ = 12) and upper ($N_e$ = 2417) bounds of the posterior distributions of $N_e$ obtained for Confluence and High Falls. Assuming a generation time of 1–3 years, and a starting $F_{ST}$ of 0 (no reproductive isolation before dam construction), it would require 1–3 years (one generation) to generate an $F_{ST}$ of 0.014 at $N_e$ = 12, and 70–210 years (70 generations) to generate an $F_{ST}$ of 0.014 at $N_e$ of 2418. These findings together suggest that the genetic differences observed between Cape Fear shiners at the Confluence and High Falls localities may be relatively recent and could have arisen from reduced gene flow caused by the dam built in 1922.

The Bayesian coalescent analysis provided evidence for a decline in effective size of one to two orders of magnitude in the two populations, suggesting that Cape Fear shiners may not have always existed in low numbers as suggested by Snelson (1971). The estimated time of the decline was within the last ~10–25 435 years with highest posterior probability of occurrence 126–2007 years ago. The lower bound of this time interval is compatible with the timing of recent anthropogenic degradation of habitat within the Cape Fear River drainage. During the past century, several dams and hydroelectric plants have been constructed in various parts of the drainage (US Fish and Wildlife Service 1988) and undoubtedly modified flow regimes and habitat quality. Increased sedimentation from soil erosion and agricultural runoff has been particularly noteworthy in the drainage (NCDWQ 2000) and would be expected to impact negatively Cape Fear shiners because of their feeding habits and spawning requirements (Howard 2003). Finally, overall water quality in the drainage has deteriorated

significantly due to relatively high levels of contaminants such as zinc, lead, DDT and chlordanes (Howard 2003).

There are caveats to the above inferences. The hierarchical model implemented in MSVAR1.3 assumes a stepwise model of mutation and an enclosed population. Mutations potentially can involve multistep changes in allele size that could create gaps in the distribution of allele sizes and produce a pattern of variation similar to that generated by a population bottleneck (Storz & Beaumont 2002). Similarly, immigration of divergent alleles from a different population could create such gaps if the immigrant alleles are outside the range of allele sizes of the recipient population (Garza & Williamson 2001). However, multistep changes appear to be very infrequent relative to single-step changes, and thus are expected to potentially affect only a small proportion of microsatellites (Storz & Beaumont 2002). Because the hierarchical model tends to give less weight in multilocus data sets to microsatellites displaying atypical allele-size distributions (Storz & Beaumont 2002), deployment of a large number of microsatellites, as in this study, is expected to recover adequately demographic signals inherent in the data. Finally, with regard to immigration, the genetic heterogeneity in both microsatellites and mtDNA detected between Confluence and High Falls indicates that migration between the two populations is very limited or absent. Even if migration is limited, allele sizes at the loci studied are sufficiently similar, such that allele-size gaps would not be expected to occur.

Both the coalescent-based analysis of Kuhner *et al*. (1998) and the mismatch distributions of pairwise differences among mtDNA haplotypes were consistent with population expansion at both localities. The estimated time since expansion based on the coalescent-based analysis was between ~68 500 and 510 000 years ago; similarly, the estimated time since expansion from mismatch-distribution analysis was between ~12 700 and 1.13 million years ago, with highest probability between 92 320 and 522 240 years ago. These times since population expansion, based on mtDNA, are two to three orders of magnitude larger than the estimated time since the population decline detected by variation in the microsatellites. This highlights the different time scales exposed by the two types of genetic markers. The estimate based on mtDNA is consistent with population expansion following a glacial cycle during the Pleistocene, whereas the population decline detected by the microsatellites dates back maximally to the last 25 000 years, and minimally within the last few decades. Similar results were obtained by Lavery *et al*. (1996), Lessios *et al*. (2001), and Koskinen *et al*. (2002), who found that genetic patterns based on mtDNA data were consistent with historical population expansion in species that today are in severe population decline. The results are also consistent with the outcome of a simulation conducted by Rogers (1996); population declines could follow expansion but not affect a genetic signature of expansion in mtDNA polymorphism.

Estimates of $N_e$ (the current effective population size) ranged from ~7 to ~4463, with highest probability of $N_e$ of 513 and 85 for Confluence and High Falls, respectively. The minimum effective size needed to ensure long-term integrity is still a matter of debate. A few hundred are, in theory, needed to maintain the equilibrium between loss of adaptive genetic variation due to genetic drift and its replacement by mutation (Franklin 1980; Schultz & Lynch 1997). However, for several reasons Lynch & Lande (1998) argued that the target effective size for a conservation programme for an endangered species should be in the range of 1000–5000. Thus, our estimates of current effective population size in Cape Fear shiner would be close to or lower than these minimum values. Our studies to date indicate that while genetic diversity appears relatively high in the populations of Cape Fear shiners at the two localities, both populations appear to have undergone significant, recent decline in effective size. In addition, effective sizes in the two populations are close to or lower than the minimum effective sizes conservatively recommended for management of endangered species (Lynch & Lande 1998). Alternatively, while genetic considerations are important relative to management of endangered species, they typically affect populations on a longer time scale than ecological and other impacts that despoil or degrade habitats. Numerous examples presented in Caro & Laurenson (1994) document the situation that human-induced perturbations typically present greater challenges to most threatened and endangered species, and without question the altered habitat in the Cape Fear drainage (Howard 2003) likely presents the greatest threat to persistence of the species. Of future concern should be whether deteriorating habitat in the Cape Fear drainage synergistically impacts genetic diversity in a negative way.

## Acknowledgements

We thank A. Howard, N. Jeffers, J. (Johnson) Ratcliffe, D. Laplante and D. Rabon for assistance in procuring specimens, F. Deal and D. Deal for providing accommodations during field work, T. Turner for providing advice regarding data analysis, M.A. Beaumont for providing us with the program MSVAR1.3 and C. Burridge for assistance in the laboratory. Comments on a draft of the manuscript were kindly provided by R. Edwards and T. Turner. Sampling was carried out under Endangered/Threatened Species Subpermittee Authorization Number SA 02–02 (US Department of the Interior). This work was supported by the US Fish and Wildlife Service (US Department of the Interior) under Agreement no. 1448-40181-00-G-119, and by the Texas Agricultural Experiment Station under Project H-6703. Views expressed in the paper are those of the authors and do not necessarily reflect views of the sponsors. The study is Contribution no. 123 of the Center for Biosystematics and Biodiversity at Texas A&M University.

## References


Akst EP, Boersma PD, Fleischer RC (2002) A comparison of genetic diversity between the Galápagos penguin and the Magellanic penguin. *Conservation Genetics*, **3**, 375–383.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 389–522.

Beaumont MA (1999) Detecting population expansion and decline using microsatellites. *Genetics*, **153**, 2013–2029.

Bermingham E, McCafferty SS, Martin AP (1997) Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. In: *Molecular Systematics of Fishes* (eds Kocher TD, Stepien CA), pp. 113–128. Academic Press, San Diego, CA.

Burridge CP, Gold JR (2003) Conservation genetic studies of the endangered Cape Fear shiner, *Notroips mekistocholas* (Teleostei: Cyprinidae). *Conservation Genetics*, **4**, 219–225.

Caro TM, Laurenson MK (1994) Ecological and genetic factors in conservation: a cautionary tale. *Science*, **263**, 485–486.

Cloutman DG, Harrell RD (1987) Life-history notes on the whitefin shiner, *Notropis niveus* (Pisces, Cyprinidae). *Copeia*, **1987**, 1037–1040.

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 461–473.

Doiron S, Bernatchez L, Blier PU (2002) A comparative mitogenomic analysis of the potential adaptative value of Arctic charr mtDNA introgression in brook charr populations (*Salvelinus fontinalis* Mitchill). *Molecular Biology and Evolution*, **19**, 1902–1909.

Farris JS, Källersjö M, Kluge AG, Bult C (1995a) Testing for significance of incongruence. *Cladistics*, **10**, 315–319.

Farris JS, Källersjö M, Kluge AG, Bult C (1995b) Constructing a significance test for incongruence. *Systematic Biology*, **44**, 570–572.

Federal Register (1987) *Federal Register*, 52, no. 186. September 25. Office of the Federal Register, Washington DC.

Frankham R (1995) Inbreeding and extinction: a threshold effect. *Conservation Biology*, **9**, 792–799.

Frankham R, Ballou JD, Broscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.

Franklin IR (1980) Evolutionary change in small populations. In: *Conservation Biology: an Evolutionary–Ecological Perspective* (eds Soulé M, Wilcox B), Chapter 8, pp. 135–149. Sinauer, Sunderland, MA.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. *Genetics*, **147**, 915–925.

Fu Y-X, Li W-H (1993) Statistical tests of neutral mutations. *Genetics*, **133**, 693–709.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.

Gold JR, Saillant E, Burridge CP, Blanchard A, Patton JC (2004) Population structure and effective size reduction in critically endangered Cape Fear shiners *Notropis mekistocholas*. *Southeastern Naturalist*, **3**, 89–102.

Goudet J (1995) FSTAT version 1.2: a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Guo S, Thompson E (1992) Performing the exact test of Hardy–Weinberg proportion from multiple alleles. *Biometrics*, **48**, 361–372.

Harrell RD, Cloutman DG (1978) Distribution and life history of sandbar shiner, *Notropis scepticus* (Pisces, Cyprinidae). *Copeia*, **1978**, 443–447.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **21**, 160–174.

Hauser L, Adcock GJ, Smith PJ, Ramirez JHB, Carvalho GR (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences USA*, **99**, 11742–11747.

Heath DD, Buch C, Kelly J, Atagi DY (2002) Temporal change in genetic structure and effective population size in steelhead trout (*Oncorhynchis mykiss*). *Molecular Ecology*, **11**, 197–214.

Hedrick PW (1999) *Genetics of Populations*, 2nd edn. Jones and Bartlett, Sudbury, MA.

Hilton-Taylor C (2000) *2000 IUCN Red List of Threatened Species*. IUCN, Cambridge, UK.

Howard AK (2003) *Influence of instream physical habitat and water quality on the survival and occurrence of the endangered Cape Fear shiner*. PhD Dissertation, North Carolina State University, Raleigh, NC.

Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **11**, 424–429.

Koskinen M, Knizhin I, Primmer CP, Schlötterer C, Weiss S (2002) Mitochondrial and nuclear DNA phylogeography of *Thymallus* spp. (grayling) provides evidence of ice-age mediated environmental perturbations in the world's oldest body of fresh water, Lake Baikal. *Molecular Ecology*, **11**, 2599–2611.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429–434.

Lance SL, Maldonado JE, Bocetti CI, Pattee OH, Ballou JD, Fleischer RC (2003) Genetic variation in natural and translocated populations of the endangered Delmarva fox squirrel (*Sciurus niger cinerus*). *Conservation Genetics*, **4**, 707–718.

Lavery S, Moritz C, Fielder DR (1996) Genetic patterns suggest exponential population growth in a declining species. *Molecular Biology and Evolution*, **13**, 1106–1113.

Lessios HA, Garrido MJ, Kessing BD (2001) Demographic history of *Diadema antillarum*, a keystone herbivore on Caribbean reefs. *Proceedings of the Royal Society of London Series B*, **268**, 2347–2353.

Lynch M, Conery J, Bürger R (1995) Mutation meltdown and the extinction of small populations. *American Naturalist*, **146**, 489–518.

Lynch M, Lande R (1998) The critical effective size for a genetically secure population. *Animal Conservation*, **1**, 70–72.

Matthews MM, Heins DC (1984) Life-history of the redfin shiner, *Notropis umbratilis* (Pisces, Cyprinidae), in Mississipi. *Copeia*, **1984**, 385–390.

Mickevich MF, Farris JS (1981) The implications of congruence in *Menidia*. *Systematic Zoology*, **30**, 351–370.

Moran P (2002) Current conservation genetics building an ecological approach to the synthesis of molecular and quantitative genetic methods. *Ecology of Freshwater Fish*, **11**, 30–55.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, NY.

North Carolina Division of Water Quality (NCDWQ) (2000) Cape Fear River basinwide water quality management plan. Raleigh, NC, 2000. In: *Influence of instream physical habitat and water quality*

*on the survival and occurrence of the endangered Cape Fear shiner* (ed. Howard AK), p. 133. PhD Dissertation. North Carolina State University, Raleigh, NC.

North Carolina Wildlife Resources Commission (1995) Cape Fear shiner (subsection). *Annual Performance Report, Nongame and Endangered Wildlife Program*, vol. IV (Biologist, Alderman JM), pp. 18–20. Project no. E-3, Segment no. 12, North Carolina Wildlife Resources Commission, Raleigh, NC.

O'Brien SJ, Martenson JS, Miththapala S *et al.* (1996) Conservation genetics of the Felidae. In: *Conservation Genetics: Case History from Nature* (eds Avise JC, Hamrick JL), pp. 50–74. Chapman & Hall, New York, NY.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Pottern GB, Huish MT (1986) *Supplement to the Status Survey of the Cape Fear Shiner Notropis mekistocholas*. US Fish and Wildlife Service Endangered Species Office, Asheville, NC.

Pottern GB, Huish MT (1987) *Second Supplement to the Status Survey of the Cape Fear Shiner Notropis mekistocholas*. US Fish and Wildlife Service Endangered Species Office, Asheville, NC.

Raymond M, Rousset F (1995) GENEPOP version 1.2: a population genetics software for exact tests and eucemenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Rogers A (1996) Population structure and modern human origins. In: *Progress in Population Genetics and Human Evolution* (eds Donnelly P, Tavare S), pp. 55–79. Springer-Verlag, New York, NY.

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–559.

Rozas J, Sánchez-Delbarrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN *Version 2000: a Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Schultz ST, Lynch M (1997) Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. *Evolution*, **51**, 1363–1371.

Sinclair EA, Costello B, Courtenay JM, Crandall KA (2002) Detecting a genetic bottleneck in Gilbert's potoroo (*Potprous gilberti*) (Marsulialia: Potoroidae), inferred from microsatellite and mitochondrial DNA sequence data. *Conservation Genetics*, **3**, 191–196.

Snelson FF (1971) *Notropis mekistocholas*, a new herbivorous cyprinid fish endemic to the Cape Fear river basin, North Carolina. *Copeia*, **1971**, 449–462.

Storz JF, Beaumont MA (2002) Testing for genetic evidence of population expansion and contraction: an empirical analysis of microsatellite DNA variation using a hierarchical Bayesian model. *Evolution*, **56**, 156–166.

Swofford DL (1998) PAUP*: *Phylogenetic Analysis Using Parsimony (*and other methods)*, version 4, 4.0b4a edn. Sinauer Associates, Sunderland MA.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619–633.

US Fish and Wildlife Service (1988) Cape Fear shiner recovery plan. US Fish and Wildlife Service, Atlanta, GA.

Vrijenhoek RC (1998) Conservation genetics of freshwater fish. *Journal of Fish Biology*, **53**, 394–412.

Waldick RC, Kraus S, Brown M, White BN (2002) Evaluating the effects of historic bottlenecks: an assessment of microsatellite variability in the endangered, North Atlantic right whale. *Molecular Ecology*, **11**, 2241–2249.

Warren ML, Burr BM, Walsh SJ *et al.* (2000) Diversity, distribution, and conservation status of the native freshwater fishes of the southern United States. *Fisheries*, **25**, 7–31.

Watterson G (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

## Appendix I

Summary statistics at 22 nuclear-encoded loci in two populations of Cape Fear shiners (*Notropis mekistocholas Locality*)

| | | Locality | | | | Locality | |
|---|---|---|---|---|---|---|---|
| Locus | | Confluence | High Falls | Locus | | Confluence | High Falls |
| Nme4F4.154 | *N* | 44 | 41 | Nme4F4.222 | *N* | 44 | 39 |
| | #A | 5 | 4 | | #A | 9 | 7 |
| | AR | 3.34 | 3.66 | | AR | 5.90 | 5.74 |
| | $H_E$ | 0.546 | 0.625 | | $H_E$ | 0.684 | 0.761 |
| | $P_{HW}$ | 0.882 | 0.847 | | $P_{HW}$ | 0.116 | 0.714 |
| | $F_{IS}$ | 0.001 | −0.015 | | $F_{IS}$ | 0.103 | 0.023 |
| Nme7B9.190 | *N* | 42 | 37 | Nme18A6.158 | *N* | 44 | 40 |
| | #A | 10 | 9 | | #A | 12 | 9 |
| | AR | 6.85 | 6.10 | | AR | 6.48 | 5.70 |
| | $H_E$ | 0.840 | 0.820 | | $H_E$ | 0.803 | 0.776 |
| | $P_{HW}$ | 0.777 | 0.637 | | $P_{HW}$ | 0.205 | 0.121 |
| | $F_{IS}$ | 0.064 | 0.110 | | $F_{IS}$ | 0.095 | 0.066 |
| Nme18C2.178 | *N* | 44 | 41 | Nme24B6.191 | *N* | 44 | 41 |
| | #A | 8 | 9 | | #A | 2 | 2 |
| | AR | 4.95 | 5.44 | | AR | 1.76 | 1.84 |
| | $H_E$ | 0.718 | 0.772 | | $H_E$ | 0.128 | 0.158 |
| | $P_{HW}$ | 0.075 | 0.616 | | $P_{HW}$ | 1.000 | 1.000 |
| | $F_{IS}$ | 0.114 | 0.147 | | $F_{IS}$ | −0.062 | −0.081 |
| Nme24B6.211 | *N* | 44 | 41 | Nme25C8.208 | *N* | 43 | 41 |
| | #A | 10 | 10 | | #A | 11 | 9 |
| | AR | 5.70 | 6.06 | | AR | 6.82 | 7.20 |
| | $H_E$ | 0.740 | 0.663 | | $H_E$ | 0.812 | 0.846 |
| | $P_{HW}$ | 0.632 | 0.281 | | $P_{HW}$ | 0.171 | 0.567 |
| | $F_{IS}$ | −0.044 | 0.080 | | $F_{IS}$ | 0.112 | −0.009 |
| Nme26G8.96 | *N* | 44 | 37 | Nme30D2.120 | *N* | 44 | 41 |
| | #A | 5 | 5 | | #A | 3 | 2 |
| | AR | 3.52 | 4.21 | | AR | 2.71 | 1.84 |
| | $H_E$ | 0.547 | 0.635 | | $H_E$ | 0.348 | 0.158 |
| | $P_{HW}$ | 0.472 | 0.682 | | $P_{HW}$ | 1.000 | 1.000 |
| | $F_{IS}$ | 0.169 | −0.192 | | $F_{IS}$ | −0.112 | −0.081 |
| Nme18A6.0 | *N* | 43 | 41 | Nme5B10.211 | *N* | 44 | 41 |
| | #A | 3 | 3 | | #A | 12 | 9 |
| | AR | 2.89 | 2.64 | | AR | 7.98 | 6.80 |
| | $H_E$ | 0.587 | 0.551 | | $H_E$ | 0.863 | 0.827 |
| | $P_{HW}$ | 0.102 | 0.365 | | $P_{HW}$ | 0.996 | 0.892 |
| | $F_{IS}$ | 0.090 | 0.203 | | $F_{IS}$ | 0.026 | −0.091 |
| Nme27F8.230 | *N*† | 44 | 41 | Nme12D8.125 | *N* | 44 | 39 |
| | #A | 8 | 8 | | #A | 11 | 9 |
| | AR | 6.07 | 6.23 | | AR | 7.47 | 6.14 |
| | $H_E$ | 0.757 | 0.821 | | $H_E$ | 0.869 | 0.781 |
| | $P_{HW}$ | 0.623 | 0.882 | | $P_{HW}$ | < 0.001* | < 0.001* |
| | $F_{IS}$ | −0.111 | −0.040 | | $F_{IS}$ | 0.372 | 0.672 |
| Nme15F2.174 | *N* | 44 | 41 | Nme30F12.208 | *N* | 42 | 40 |
| | #A | 6 | 7 | | #A | 14 | 14 |
| | AR | 4.31 | 5.20 | | AR | 7.66 | 7.97 |
| | $H_E$ | 0.560 | 0.705 | | $H_E$ | 0.846 | 0.874 |
| | $F_{IS}$ | −0.055 | 0.032 | | $F_{IS}$ | 0.184 | 0.256 |
| | $P_{HW}$ | 0.962 | 0.547 | | $P_{HW}$ | 0.017 | 0.002 |
| Nme33B6.125 | *N* | 44 | 41 | Nme2B10.166 | *N* | 43 | 38 |
| | #A | 5 | 7 | | #A | 7 | 6 |
| | AR | 4.37 | 5.51 | | AR | 4.88 | 4.76 |
| | $H_E$ | 0.666 | 0.738 | | $H_E$ | 0.727 | 0.651 |
| | $P_{HW}$ | 0.705 | 0.132 | | $P_{HW}$ | 0.888 | 0.754 |
| | $F_{IS}$ | −0.058 | 0.173 | | $F_{IS}$ | 0.008 | −0.010 |

**Appendix I** *Continued*

| Locus | | Locality | | Locus | | Locality | |
|---|---|---|---|---|---|---|---|
| | | Confluence | High Falls | | | Confluence | High Falls |
| Nme2B10.232 | *N* | 44 | 41 | Nme5G4.175 | *N* | 44 | 41 |
| | #A | 13 | 11 | | #A | 13 | 10 |
| | AR | 8.40 | 7.83 | | AR | 7.50 | 6.11 |
| | $H_E$ | 0.843 | 0.871 | | $H_E$ | 0.812 | 0.734 |
| | $P_{HW}$ | 0.050 | 0.705 | | $P_{HW}$ | 0.026 | 0.476 |
| | $F_{IS}$ | 0.084 | 0.076 | | $F_{IS}$ | 0.132 | −0.030 |
| Nme6A7.93 | *N* | 44 | 41 | Nme2D5.117 | *N* | 44 | 41 |
| | #A | 11 | 10 | | #A | 11 | 11 |
| | AR | 7.51 | 7.30 | | AR | 6.65 | 8.07 |
| | $H_E$ | 0.850 | 0.848 | | $H_E$ | 0.824 | 0.877 |
| | $P_{HW}$ | 0.790 | 0.472 | | $P_{HW}$ | 0.747 | 0.869 |
| | $F_{IS}$ | 0.010 | 0.049 | | $F_{IS}$ | −0.130 | 0.029 |

*Significant after Bonferroni correction.
*N* = sample size, †A = number of alleles; AR = allelic richness, $H_E$ = gene diversity, $P_{HW}$ = probability of conformity to Hardy–Weinberg genotypic expectations and $F_{IS}$ = inbreeding coefficient.