Available online at www.sciencedirect.com
SCIENCE DIRECT®
Forensic Science International 156 (2006) 9–15
Forensic Science International
www.elsevier.com/locate/forsciint

# Application of hypervariable genetic markers to forensic identification of 'wild' from hatchery-raised red drum, *Sciaenops ocellatus*

M.A. Renshaw [a], E. Saillant [a], R.E. Broughton [b], J.R. Gold [a,*]

[a] *Center for Biosystematics and Biodiversity, Texas A&M University, College Station, Texas, TX 77843-2258, USA*
[b] *Oklahoma Biological Survey and Department of Zoology, University of Oklahoma, 111 E. Chesapeake Street, Norman, Oklahoma, OK 73019, USA*

Received 31 December 2004; received in revised form 2 May 2005; accepted 3 May 2005

**Abstract**

Forensic identification of 'wild' versus hatchery-produced (cultured) red drum (*Sciaenops ocellatus*), an economically important marine fish in the southern United States, was assessed using hypervariable nuclear-encoded microsatellites and sequences of mitochondrial DNA. Both genotype exclusion and likelihood-ratio tests successfully identified 'wild' and 'cultured' individuals within requisite error bounds and within the context of complete parental sampling. Of the two, genotype exclusion was more effective, producing satisfactory results with fewer microsatellites and larger allowable error rates. Assignment tests proved ineffective, most likely because of the low level of genetic divergence between the sampled populations. An optimal, minimum set of ten markers that will reduce potential genotyping costs is identified. Results of the study should allay concerns regarding identification of 'wild'-caught fish sold illegally.


*Keywords:* Microsatellites; Forensic genetics; 'Wild' versus hatchery-raised fish

The existence of a legal market for wildlife meat and/or parts raises the issue of eventual sale of a legal species obtained or marketed illegally. Examples include sale without proper licensing, use of illegal gear for capture, or procurement from a protected area such as a park or wildlife reserve [1–3]. A forensic issue in these cases is how to discriminate the illegal samples from the legal ones. In this paper, we focus on this issue using as an example the red drum, *Sciaenops ocellatus*, an estuarine-dependent sciaenid fish found along the Atlantic and Gulf coasts of the southeastern United States. Juvenile red drum are exposed to heavy recreational fishing in bays and estuaries, and declines in population abundance led to rather extreme fishing restrictions in most U.S. waters [4]. One consequence is that red drum are now cultured in several southern states [5–7], and in Texas there currently is an extensive stock-enhancement program where red drum fingerlings are released into numerous Texas bays and estuaries [5].

The ready availability of hatchery-spawned red drum fingerlings and juveniles makes them ideal for offshore aquaculture. The creation of a market for red drum cultured offshore, however, could generate illegal sale of 'wild' red drum if 'wild' individuals (caught by angling or netting) are marketed as hatchery-produced individuals. We evaluated the potential for nuclear-encoded microsatellites and sequences from the mitochondrially encoded D-loop or control region to distinguish 'wild'-caught red drum from hatchery-produced individuals that could be grown

* Corresponding author. Tel.: +1 979 847 8778; fax: +1 979 845 4096.
*E-mail address:* goldfish@tamu.edu (J.R. Gold).

0379-0738/$ – see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.forsciint.2005.05.038

in offshore aquaculture facilities. The study was part of a large, multi-disciplinary project (http://www.masgc.org/oac) designed to assess the potential of offshore aquaculture in the northern Gulf of Mexico. The project design was to grow red drum obtained from hatcheries in Texas in cage facilities located offshore of Ocean Springs, Mississippi. We used a sample of 102 'wild' red drum caught in Biloxi Bay, Mississippi, to represent the 'wild' population; 'hatchery' red drum were represented by (simulated) offspring from a sample of 45 individuals from the broodstock used by the Texas Parks and Wildlife Department (TPWD) in red drum stock enhancement.

We employed three different statistical approaches to distinguish 'wild' from hatchery-produced fish: (i) genotype exclusion and match probabilities [8-9]; (ii) likelihood-ratio tests [10]; and (iii) assignment tests [11,12]. The first two approaches require knowledge of at least one parental genotype and use these genotypes with those of putative offspring to identify parent-offspring relationships. Genotype exclusion relies on direct scoring of matching and non-matching alleles at individual markers and employs match probabilities to evaluate the probability of exclusion outcomes. Likelihood-ratio tests compare the likelihood that genotypes of putative parents and offspring were generated by a parent-offspring relationship to the likelihood that these genotypes occur in unrelated individuals. Assignment tests evaluate exclusion probability with regard to the parental population as a whole and thus do not test specific parent-offspring relationships. Assignment tests only require knowledge of allele distributions in the population from which parents were drawn. The relative power of the three approaches to distinguish 'wild' from cultured fish was assessed.

## 1. Materials and methods

### 1.1. Samples

Clips (approximately 2–3 $cm^3$) from the upper lobe of the caudal fin were taken from red drum broodfish held in each of nine brood tanks at the CCA/CPL Marine Development Center in Corpus Christi, Texas. Each brood tank contained two sires and three dams (45 fish total). Fin clips were fixed in 95% ethanol and stored at room temperature. Heart tissues, frozen in liquid nitrogen, from a total of 102 'wild' red drum sampled from Biloxi Bay, Mississippi, were kindly provided by J. Franks of the Gulf Coast Research Laboratory in Ocean Springs, Mississippi. DNA from all individuals was isolated and purified using methods [13] routinely employed in our laboratory.

### 1.2. Polymerase chain reaction

PCR reactions for 25 microsatellite markers were performed in 10 μl volumes containing 1 μl (100 ng) of DNA, 1 μl of 10× reaction buffer (500 mM KCl, 200 mM Tris–HCl [pH 8.4]), 1.5 mM $MgCl_2$, 2.5 mM of each dNTP, 5 pmols of each primer, and 0.5 units *Taq* polymerase (GibcoBRL, Carlsbad, California). The PCR protocol for all microsatellites consisted of initial denaturation at 94 °C for 3 min, followed by 35 cycles consisting of 30 s at 94 °C, 45 s at the optimized annealing temperature, extension for 1 min at 72 °C, and a final extension of 10 min at 72 °C. PCR primer sequences, repeat motif, and optimal annealing temperature for each microsatellite are given in [14].

A fragment of 370 base pairs from the D-loop or control region of the red drum mitochondrial (mt)DNA was amplified in 50 μl reactions. Each reaction contained 5 μl of 10× reaction buffer (500 mM KCl, 100 mM Tris–HCl [pH 8.5]), 1.5 mM $MgCl_2$, 200 μM of each dNTP, 0.5 μM of each primer, 2.5 units of *Taq* polymerase, approximately 100 ng of template DNA, and ultrapure water. The PCR protocol consisted of an initial denaturation at 94 °C for 30 s, followed by 30 cycles of denaturation at 94 °C for 10 s, annealing at 55 °C for 15 s, and extension for 45 s at 72 °C. Amplification and sequencing primers are given in [15].

### 1.3. Basic statistics and parental pair exclusion probability

Genotypes for 25 nuclear-encoded microsatellites and sequences from the mtDNA control region were acquired from 144–147 fish (45 broodfish from the TPWD hatchery, hereafter *Broodfish*, and 99–102 'wild' fish from Biloxi Bay, hereafter *Biloxi Bay*). Exact tests of conformity of genotypes at each microsatellite marker to Hardy–Weinberg equilibrium expectations were implemented using GENEPOP 3.3 [16]. Allele number and gene diversity (expected heterozygosity or $H_E$) for each microsatellite over the entire data set were obtained using CERVUS [17]. Microsatellites were ranked based on their estimated gene diversity (1–25, high $H_E$ to low $H_E$). Genetic variability for the mtDNA marker was measured by number of haplotypes and nucleon diversity (probability that two individuals will differ in mtDNA haplotype), using the program DNASP [18].

The parental-pair exclusion probability for each microsatellite was obtained using FAMOZ [19]. This probability reflects the effectiveness of a given set of microsatellites to identify a parent-offspring relationship [20]. Here, it represents the probability of a given set of microsatellites to exclude 'false' parent pairs (sires and dams from *Broodfish*) for a given individual 'wild' fish from *Biloxi Bay*. Exclusion probabilities were calculated for individual microsatellites and various subsets of microsatellites.

#### 1.3.1. Genotype exclusion and match probabilities

The genotype-exclusion approach outlined in [8] was used to compare genotypes from the 45 *Broodfish* (18 sires and 27 dams) with genotypes from the 102 *Biloxi Bay* fish. Genotypes from the 45 *Broodfish* also were compared with 20 simulated '*Broodfish* progeny' genotypes for each of the

possible 486 *Broodfish* combinations (27 dams × 18 sires − 9720 genotypes total) generated using PROBMAX-G [21]. For nuclear-encoded markers (microsatellites), genotype comparisons and exclusion of incompatible individuals from *Biloxi Bay* followed Mendelian principles (i.e., for each microsatellite, offspring must have one allele derived from one of the 27 dams and one allele derived from one of the 18 sires in *Broodfish*). Fish from *Biloxi Bay* or from the simulated *Broodfish* progeny that failed to meet this criterion were excluded as possible offspring from sire x dam combinations in *Broodfish*. PROBMAX-2 [21] was used to perform exclusions. The minimum number of microsatellites needed for 100% exclusion was assessed by carrying out exclusion 'runs' using subsets of 5, 10, 15, 20, and 25 microsatellites. Selection of microsatellites in subsets was based on 'polymorphism' rank, i.e., the subset of five microsatellites was composed of the markers ranked 1–5 (the five highest gene diversities), the subset of ten markers was composed of those ranked 1–10 (the 10 highest gene diversities), and so forth. Each exclusion analysis also was run with mismatches allowed at three, two, one and zero microsatellites in order to evaluate potential impact of mutations or genotyping errors caused by null alleles, and/or mis-scoring [22,23]. Mismatches involved randomly any of the *n* microsatellites in a subset and could occur at either one or both alleles at a given microsatellite. For the mitochondrial DNA marker, progeny mtDNA haplotypes were required to match a haplotype from one of the *Broodfish* dams. *Biloxi Bay* fish that failed to meet this requirement were excluded.

Match probabilities were calculated based on the most common, *expected* hatchery genotype (determined from allele frequencies in *Broodfish*) and the *expected* frequency of this most-common hatchery genotype among the 'wild' population (determined from allele frequencies in *Biloxi Bay*). The match probability for the mtDNA marker was estimated as [*x*/99], where *x* was the number of *Biloxi Bay* fish possessing the most common female *Broodfish* mtDNA haplotype divided by the total number (99) of assayed *Biloxi Bay* fish. Estimates were computed per genetic marker (microsatellites and mtDNA) and multiplied across all markers to provide multi-locus estimates for the subsets of 5, 10, 15, 20, and 25 microsatellites used in exclusion analyses. An adjusted probability that provided a minimum threshold for match probabilities was generated to determine the number of markers necessary for credible exclusion. The adjusted probability was estimated as ...

$$p_x \leq 1 - (1 - \alpha)^{1/N}$$

where $p_x$ is the adjusted threshold probability; $\alpha$, the defined error rate; and $N$, the estimated number of individuals in the candidate source population [24]. For this study, the threshold probability was estimated using an error rate ($\alpha$) of 99% [24]; the estimated number of individuals ($N$) in the population was $7.7 \times 10^6$, the upper 90% confidence interval of an estimate of census size for red drum in the northern Gulf of Mexico [25,26].

#### 1.3.2. Likelihood-ratio tests

Likelihood-ratio tests, as implemented in FAMOZ [19,27], were used to assign parentage in both single-parent (incomplete parental sampling) and parent-pair (complete parental sampling) contexts. Likelihood ratios (LOD scores) compare the likelihood of an individual (or two individuals) being the parent (or parent pair) of a given offspring versus the likelihood that they are not (the parent or parent pair) (27). The simulation component in FAMOZ was used to randomly sample gametes from genotyped (*Broodfish*) parents and combine them to produce a user-specified number (10,000 for this study) of offspring genotypes. These genotypes were then used to generate a distribution of LOD scores; the lower bound of the simulated LOD-score distribution was employed as the critical threshold for parentage assignment. The variable representing the error rate for LOD score calculation [19,27] was manipulated in FAMOZ to evaluate potential impacts of typing errors. For parentage assignment, LOD scores were generated for both single parent-offspring and parent pair-offspring scenarios, evaluating all pairwise comparisons of candidate parents (45 *Broodfish*) versus the 102 candidate offspring from *Biloxi Bay* and the 9720 simulated *Broodfish* progeny (generated by PROBMAX-G). LOD scores that exceeded the appropriate threshold (estimated from FAMOZ simulations) were indicative of parent-offspring relationships between *Broodfish* and *Biloxi Bay* or between *Broodfish* and the simulated *Broodfish* progeny. Likelihood analyses were run with subsets of 5, 10, 15, 20, and 25 microsatellites, with and without the mtDNA marker. LOD calculation error rates (typing errors) of 0.1%, 1%, and 5% were simulated for each data set.

#### 1.3.3. Assignment tests

A frequency-based assignment test [11], as implemented in GENECLASS-2 [28], was used to estimate the probability that the 102 *Biloxi Bay* fish and 9720 simulated *Broodfish* progeny could be assigned to either the 'wild' reference population, derived from the 102 *Biloxi Bay* genotypes, or the cultured reference population, derived from the 45 *Broodfish* genotypes. A Monte-Carlo resampling algorithm [29] was used to generate statistical thresholds for assignment, using Type I error rates of 5%, 1%, and 0.1%. A total of 20,000 individual genotypes (10,000 per reference population) were generated with the Monte-Carlo resampling method by drawing, with replacement, multilocus gametes from randomly chosen individuals in both reference populations. The frequency-based assignment test was employed with the subsets of 5, 10, 15, 20, and 25 microsatellites.

## 2. Results

Genotypes and allele frequencies for all 25 microsatellites and for mtDNA in both *Broodfish* and *Biloxi Bay* may be

found in Appendix 1 and Appendix 2 (http://wfsc.tamu.edu/docs). Tests of conformity of microsatellite genotypes to Hardy–Weinberg equilibrium expectations were non-significant in both samples, following sequential Bonferroni correction [30]. Summary statistics for the 25 microsatellites over all individuals assayed may be found in Appendix 3 (http://wfsc.tamu.edu/docs). The number of alleles per microsatellite ranged from five (*Soc*156, *Soc*444) to 33 (*Soc*428); gene diversity ranged from 0.541 (*Soc*444) to 0.955 (*Soc*428). These values are well within ranges observed for 524 microsatellites from 78 species of fishes [31]. Polymorphism rank (1–25), parental pair exclusion probability, and match probability for each microsatellite also are listed in Appendix 3. Parental pair exclusion probability ranged from 0.408 (*Soc*444) to 0.983 (*Soc*428), while match probability ranged from 0.404 (*Soc*60) to 0.017 (*Soc*428). A total of 90 mtDNA haplotypes were found over all 144 individuals assayed. Individual haplotypes and their frequencies may be found in Appendix 1 and Appendix 4 (http://wfsc.tamu.edu/docs). Nucleon diversity was 0.987 and within the range (0.95–1.00) found previously [15] for red drum. The estimated match probability for mtDNA was 0.333.

Results for genotype exclusion analyses, using various subsets of the 25 available microsatellites, are given in Table 1. At least 99% of the 'wild' fish (100%, in most instances) were successfully excluded when 10 or more microsatellites were used (results using 20 and 25 microsatellites not shown). When only five microsatellites were used, exclusion success was reduced when mismatches were allowed at more than one microsatellite (down to 13% exclusion when three mismatches were allowed). Addition of the mtDNA marker considerably strengthened exclusion analysis. When considering just the subset of five microsatellites plus mtDNA, 97% of the 'wild' samples were excluded when mismatches were allowed at three microsatellites; whereas only 13% of the 'wild' samples were excluded when mtDNA was not included (Table 1). As anticipated, exclusion analysis failed to exclude any of the 9720 simulated *Broodfish* progeny genotypes in all analyses performed (Table 1).

Table 1
Results of genotype exclusion: numbers represent percentage of individuals from sampling groups excluded as offspring from 18 sires and 27 dams in *Broodfish*

| Input markers | Mismatches allowed[a] | 'Wild' | | | 'Simulated' | | |
|---|---|---|---|---|---|---|---|
| | | 5 | 10 | 15 | 5 | 10 | 15 |
| Microsatellites only | 3 | 13 | 99 | 100 | 0 | 0 | 0 |
| | 2 | 77 | 100 | 100 | 0 | 0 | 0 |
| | 1 | 99 | 100 | 100 | 0 | 0 | 0 |
| | 0 | 100 | 100 | 100 | 0 | 0 | 0 |
| Microsatellites and mtDNA | 3 | 97 | 100 | 100 | 0 | 0 | 0 |
| | 2 | 100 | 100 | 100 | 0 | 0 | 0 |
| | 1 | 100 | 100 | 100 | 0 | 0 | 0 |
| | 0 | 100 | 100 | 100 | 0 | 0 | 0 |

Sampling groups were 102 individuals from *Biloxi Bay* ('Wild') and 9720 simulated *Broodfish* progeny ('Simulated'). Runs incorporated 5, 10, and 15 microsatellites (details in text).

[a] Mismatches allowed at 3, 2, 1, and 0 microsatellites.

Multi-locus match probabilities for subsets of 5, 10, 15, 20, and 25 microsatellites, including and excluding the mitochondrial marker, are given in Table 2. Subsets with more than five microsatellites easily surpassed the adjusted minimal threshold probability of $1.31 \times 10^{-9}$, estimated using an error rate ($\alpha$) of 99%, whether or not the mitochondrial marker was included. The similarity in multi-locus match probabilities both with and without the mitochondrial marker was anticipated as the match probability (0.333) for the mitochondrial marker was relatively non-specific. The multi-locus match probabilities are comparable to the results from exclusion analysis; a subset of ten microsatellites was sufficient to exclude all 'wild' individuals even allowing for mismatches at up to three microsatellites (genotype exclusion) and to potentially assert unique source attribution (match probabilities).

Results of likelihood analyses (Table 3) indicated that 15 microsatellites were necessary to exclude all *Biloxi Bay* individuals and include all simulated *Broodfish* progeny as having been produced from *Broodfish* sires and dams at a typing error rate of 0.1% and with complete parental sampling. The analysis was less effective even for the larger sets of 20 and 25 microsatellites when error rates of 1% and 5% were evaluated (percentages for 25 microsatellites were essentially the same as for 20 microsatellites). Inclusion of the mitochondrial marker increased the overall effectiveness of the analysis and reduced the number of microsatellites needed for complete exclusion of *Biloxi Bay* fish (and complete inclusion of 'simulated' progeny) to 10 (Table 3). The likelihood approach was more effective with incomplete parental sampling (genotypes of only one parent are known) than with complete parental sampling (genotypes of both parents known) for assigning simulated *Broodfish* progeny to *Broodfish*. Complete parental sampling was more effective for excluding *Biloxi Bay* individuals as possible offspring from *Broodfish* (Table 3).

Table 2
Match probabilities for subsets of 5, 10, 15, 20, and 25 microsatellite markers with and without the mitochondrial (mtDNA) marker

| Number of microsatellites | Match probabilities (with mtDNA) | Match probabilities (without mtDNA) |
|---|---|---|
| 5 | $1.43 \times 10^{-8}$ | $4.29 \times 10^{-8}$ |
| 10 | $1.79 \times 10^{-14a}$ | $5.39 \times 10^{-14a}$ |
| 15 | $7.99 \times 10^{-19a}$ | $2.40 \times 10^{-18a}$ |
| 20 | $5.66 \times 10^{-22a}$ | $1.70 \times 10^{-21a}$ |
| 25 | $2.67 \times 10^{-24a}$ | $8.03 \times 10^{-24a}$ |

[a] Match probabilities that exceed the threshold probability estimated using an error rate ($\alpha$) of 99%.

Table 3
Results of the likelihood-ratio tests: numbers represent percentage of individuals excluded as offspring from 18 sires and 27 dams in *Broodfish*[a]

| Parental sampling | Input markers | Typing error (%) | 'Wild' | | | | 'Simulated' | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | 5 | 10 | 15 | 20 | 5 | 10 | 15 | 20 |
| Complete | Microsatellites | 0.1 | 100 | 100 | 100 | 100 | 7 | 1 | 0 | 0 |
| | | 1.0 | 69 | 100 | 100 | 100 | 36 | 19 | 7 | 9 |
| | | 5.0 | 35 | 56 | 68 | 75 | 38 | 31 | 24 | 38 |
| | Microsatellites and mtDNA | 0.1 | 100 | 100 | 100 | 100 | 4 | 0 | 0 | 0 |
| | | 1.0 | 98 | 100 | 100 | 100 | 17 | 9 | 3 | 5 |
| | | 5.0 | 76 | 76 | 94 | 91 | 3 | 10 | 11 | 24 |
| Incomplete | Microsatellites | 0.1 | 0 | 3 | 8 | 5 | 0 | 0 | 0 | 0 |
| | | 1.0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | 5.0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Microsatellites and mtDNA | 0.1 | 0 | 6 | 9 | 5 | 0 | 0 | 0 | 0 |
| | | 1.0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | 5.0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Sample groups are 102 individuals from *Biloxi Bay* ('Wild') and 9720 simulated *Broodfish* progeny ('Simulated') Parental sampling was complete (both parents) or incomplete (one parent), using 5, 10, 15, and 20 microsatellites.

[a] Typing error defined as the rate at which the true genotype at a particular locus in an individual is replaced by a random incorrect genotype.

The frequency-based assignment test was the least effective of the three approaches employed. With all 25 microsatellites, only 6% of the 102 *Biloxi Bay* fish and 45% of the simulated *Broodfish* progeny were assigned to their correct group at a statistical threshold ($\alpha$) level of 0.01 (Table 4). In addition, the percentages at a threshold of 0.05 were marginally worse (2%) for the *Biloxi Bay* fish and equivalent (45%) for the simulated *Broodfish* progeny. Even when assignments were made solely on the basis of the most probable reference population (no statistical threshold applied), only 35% of the *Biloxi Bay* fish were correctly assigned. A Bayesian-based assignment test [12] performed essentially the same (data not shown).

Table 4
Results of frequency-based assignment tests: numbers represent percentage of individuals assigned to their appropriate sample group

| Statistical threshold[a] | 'Wild' | | | | | 'Simulated' | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 10 | 15 | 20 | 25 | 5 | 10 | 15 | 20 | 25 |
| 0.001 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 17 |
| 0.01 | 1 | 3 | 3 | 1 | 6 | 3 | 3 | 3 | 3 | 45 |
| 0.05 | 4 | 10 | 8 | 11 | 2 | 13 | 13 | 13 | 16 | 48 |
| None | 73 | 73 | 77 | 69 | 35 | 88 | 94 | 96 | 98 | 100 |

Sample groups were 102 individuals from *Biloxi Bay* ('Wild') and 9720 simulated *Broodfish* progeny ('Simulated') Runs incorporated 5, 10, 15, 20, and 25 microsatellites.

[a] Statistical threshold represents Type I error rates ($\alpha$) of 0.001, 0.01, and 0.05. None refers to analyses run with no specified $\alpha$ threshold, i.e., where individuals were assigned to the highest probability reference population without additional statistical requirements.

## 3. Discussion

Our study involved forensic application of microsatellite genotyping to potential legal issues regarding a popular, recreationally harvested species of fish found along the Atlantic and Gulf coasts of the United States. Whereas non-forensic application of microsatellite genotyping (e.g., stock-structure analysis) typically allows for an error rate $\leq$5% [32], forensic application generally mandates a confidence level of $\geq$99% [24]. Of the three statistical tests assessed in our study, genotype exclusion and likelihood-ratio tests successfully identified 'wild' and 'cultured' individuals within requisite error bounds. Assignment tests were unsuccessful within any measure of confidence.

Of the two successful approaches, genotype exclusion was more effective than likelihood-ratio tests, producing satisfactory results with input from fewer microsatellite markers (10 versus 15) and larger allowable error rates (20% versus 0.1%). Genotype exclusion traditionally has been the preferred statistical approach given complete sampling of all candidate parents and sufficient genotypic information to exclude all (or all but one) parental pairs [22,33]. A potential weakness of genotype exclusion can be an increase in false exclusions attributable to mutations or genotyping errors (i.e., null alleles, human error) [22]. Allowing for mismatches at a number of microsatellites mitigates the problems caused by genotyping errors. In our study, genotype exclusion was 100% accurate for both excluding *Biloxi Bay* fish and *not* excluding simulated *Broodfish* progeny when mismatches were allowed at two (out of ten) microsatellites. Allowing mismatches at a third microsatellite only slightly reduced accuracy (99%) for excluding *Biloxi Bay* individuals, while not reducing the accuracy for including the simulated *Broodfish* progeny.

In the case of incomplete sampling of candidate parents, the two approaches commonly used are likelihood-ratio and assignment tests [22,33]. Neither method was effective at excluding all *Biloxi Bay* individuals and *not* excluding simulated *Broodfish* progeny. The inability of likelihood-ratio tests to exclude *Biloxi Bay* individuals with incomplete parental sampling was unexpected given (i) the success of exclusion with complete parental sampling, and (ii) the success of the approach in a study of white sturgeon [34]. However, a similar contrast in outcomes between single-parent and parent-pair analysis was reported previously [35] in a study of armadillos; in that case, the most likely sire and dam were not in agreement between the two parental-sampling schemes in 79% and 49% of the cases, respectively. An explanation for the apparent discrepancy is that a candidate offspring carrying a rare or infrequent allele in a heterozygous condition may display a high single-parent LOD score when a putative parent possesses the rare allele, but may not necessarily display a high parent-pair LOD score depending on candidate genotypes for the second parent [35,36]. It also has been demonstrated [36] that with multiple loci, almost all offspring possess a sufficient number of infrequent or rare alleles to generate a negative correlation between maternal and paternal LOD scores.

Exclusion via assignment tests proved ineffective in our study. This result is most likely due to the low levels of divergence between the sampled populations. Among-group variation of red drum in the northern Gulf of Mexico [13,37] is substantially less than the $F_{ST}$ range of $\geq$0.05 recommended [38] for effective application of assignment tests. In addition, based on simulation analysis, application of assignment tests with a forensic requirement of $\alpha = 0.1\%$ would require >20 microsatellites, even with substantial variation ($F_{ST} \approx 0.1$) between or among samples.

The intuitive response for analyzing an unknown sample in a case such as in our study most likely would be to obtain genotypes from as many genetic markers as possible, attempting – by brute force – to eliminate any possibility of not excluding 'wild' individuals. Such an approach is not necessarily cost effective, incurring expenses for additional markers that may not be needed. This conundrum requires a gauge for the amount of information necessary, short of an unnecessary overstatement of the strength of evidence. Calculating match probabilities and estimating adjusted match-probability thresholds can provide such a gauge. In our study, match probabilities indicated that ten microsatellites were sufficient for forensic application, and that the additional time and effort necessary to obtain mtDNA sequences was not essential. In general, it may not be cost effective to include mtDNA sequencing when sufficient numbers of hypervariable microsatellites are available.

## Acknowledgements

Work was supported by a grant from the National Sea Grant Program (OAR-NOAA Award NA16RG1631) and by the Texas Agricultural Experiment Station (Project H-6703). We thank R. Vega and staff at the Texas Parks and Wildlife CCA/CPL Marine Development Center in Corpus Christi, Texas, and J. Franks of the Gulf Coast Research Laboratory in Ocean Springs, Mississippi, for providing samples used in the study. We also thank S. C. Bradfield for technical assistance. This paper is number 44 in the series 'Genetic Studies in Marine Fishes' and Contribution No. 134 of the Center for Biosystematics and Biodiversity at Texas A&M University.

## References

[1] C.R. Primmer, M.T. Koskinen, J. Piironen, The one that did not get away: individual assignment using microsatellite data detects a case of fishing competition fraud, Proc. R. Soc. Lond. B 267 (2000) 1699–1704.
[2] E. White, J. Hunter, C. Dubetz, R. Brost, A. Bratton, S. Edes, R. Sahota, Microsatellite markers for individual tree genotyping: application in forest crime prosecutions, J. Chem. Technol. Biotechnol. 75 (2000) 923–926.
[3] A. Blom, J. Yamindou, H.H.T. Prins, Status of the protected areas of the Central African Republic, Biol. Conserv. 118 (2004) 479–487.
[4] Gulf States Marine Fishery Commission, A summary of marine fishing laws and regulations for the Gulf States, 2002, available at: http://gsmfc.org.
[5] L.W. McEachron, C.E. McCarty, R.R. Vega, Beneficial uses of marine fish hatcheries: enhancement of red drum in Texas coastal waters, Am. Fish. Soc. Symp. 15 (1995) 161–166.
[6] T.N. Bert, R.H. McMichael Jr., R.P. Cody, A.B. Forstchen, W.G. Halstead, J. O'Hop, J.M. Ransier, M.D. Tringali, B.L. Winner, F.S. Kennedy, K.M. Leber, C.L. Neidig, Evaluating stock enhancement strategies: a multi-disciplinary approach, UJNR Tech. Rep. 30 (2001) 105–126.
[7] T.I.J. Smith, W.E. Jenkins, M.R. Denson, Overview of an experimental stock enhancement program for red drum in South Carolina, Bull. Natl. Res. Inst. Aquacult. Suppl. 3 (1997) 109–115.
[8] R. Perez-Enriquez, N. Taniguchi, Use of microsatellite DNA as genetic tags for the assessment of a stock enhancement program of red sea bream, Fish. Sci. 65 (1999) 374–379.
[9] R. Chakraborty, D.N. Stivers, B. Su, Y. Zhong, B. Budowle, The utility of short tandem repeat loci beyond human identification: Implications for development of new DNA typing systems, Electrophoresis 20 (1999) 1682–1696.
[10] S. Gerber, S. Mariette, R. Streiff, C. Bodenes, A. Kremer, Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis, Mol. Ecol. 9 (2000) 1037–1048.
[11] D. Paetkau, W. Calvert, I. Stirling, C. Strobeck, Microsatellite analysis of population structure in Canadian polar bears, Mol. Ecol. 4 (1995) 347–354.
[12] B. Rannala, J.L. Mountain, Detecting immigration by using multilocus genotypes, Proc. Natl. Acad. Sci. U.S.A. 94 (1997) 9197–9201.
[13] J.R. Gold, L.R. Richardson, Genetic studies in marine fish. IV. An analysis of population structure in the red drum (*Sciaenops ocellatus*) using mitochondrial DNA, Fish. Res. 12 (1991) 213–241.

[13] J.R. Gold, L.R. Richardson, Genetic studies in marine fish. IV. An analysis of population structure in the red drum (*Sciaenops ocellatus*) using mitochondrial DNA, Fish. Res. 12 (1991) 213–241.
[14] E. Saillant, K. Czidziel, K.G. O'Malley, T.F. Turner, C.L. Pruett, J.R. Gold, Microsatellite markers for red drum, *Sciaenops ocellatus*, Gulf Mexico Sci. 1 (2004) 101–107.
[15] S. Seyoum, M.D. Tringali, T.M. Bert, D. McElroy, R. Stokes, An analysis of genetic population structure in red drum, *Sciaenops ocellatus*, based on mtDNA control region sequences, Fish. Bull. 98 (2000) 127–138.
[16] M. Raymond, F. Rousset, An exact test for population differentiation, Evolution 49 (1995) 1280–1283.
[17] T.C. Marshall, J. Slate, L.E.B. Kruuk, J.M. Pemberton, Statistical confidence for likelihood-based paternity inference in natural populations, Mol. Ecol. 7 (1998) 639–655.
[18] J. Rozas, J.C. Sanchez-DelBarrio, X. Messeguer, R. Rozas, DnaSP, DNA polymorphism analyses by the coalescent and other methods, Bioinformatics 19 (2003) 2496–2497.
[19] S. Gerber, P. Chabrier, A. Kremer, FAMOZ: a software for parentage analysis using dominant, codominant and uniparentally inherited markers, Mol. Ecol. Notes 3 (2003) 479–481.
[20] A. Jamieson, St. C.S. Taylor, Comparisons of three probability formulae for parentage exclusion, Anim. Gen. 28 (1997) 397–400.
[21] R.G. Danzmann, PROBMAX: a computer program for assigning unknown parentage in pedigree analysis from known genotypic pools of parents and progeny, J. Hered. 88 (1997) 333.
[22] A.G. Jones, W.R. Arden, Methods of parentage analysis in natural populations, Mol. Ecol. 12 (2003) 2511–2523.
[23] J.M. Pemberton, J. Slate, D.R. Bancroft, J.A. Barrett, Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies, Mol. Ecol. 4 (1995) 249–252.
[24] B. Budowle, R. Chakraborty, G. Carmody, K.L. Monson, Source attribution of a forensic DNA profile, Forensic Sci. Commun. 2 (2000) 3.
[25] S. Nichols, An estimate of the size of red drum spawning stock using mark/recapture, National Marine Fisheries Service, Southeast Fisheries Science Center, Mississippi Laboratory, Pascagoula, Mississippi, 1988.
[26] K.M. Mitchell, T. Henwood, Red drum (*Sciaenops ocellatus*) tag/recapture experiment, National Marine Fisheries Service, Southeast Fisheries Science Center, Mississippi Laboratory, Pascagoula, Mississippi, 1999.
[27] S. Gerber, S. Mariette, R. Streiff, C. Bodenes, A. Kremer, Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis, Mol. Ecol. 9 (2000) 1037–1048.
[28] S. Piry, A. Alapetite, J.M. Cornuet, D. Paetkau, L. Baudouin, A. Estoup, GeneClass2: a software for genetic assignment and first generation migrant detection, J. Hered. 95 (2004) 536–539.
[29] D. Paetkau, R. Slade, M. Burden, A. Estoup, Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power, Mol. Ecol. 13 (2004) 55–65.
[30] W.R. Rice, Analyzing tables of statistical tests, Evolution 43 (1989) 223–225.
[31] J.A. DeWoody, J.C. Avise, Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals, J. Fish Biol. 56 (2000) 461–473.
[32] F.W. Allendorf, S.R. Phelps, Use of allelic frequencies to describe population structure, Can. J. Fish. Aquat. Sci. 38 (1981) 1507–1514.
[33] A.J. Wilson, M.M. Ferguson, Molecular pedigree analysis in natural populations of fishes: approaches, applications, and practical considerations, Can. J. Fish. Aquat. Sci. 59 (2002) 1696–1707.
[34] J.A. Rodzen, T.R. Famula, B. May, Estimation of parentage and relatedness in the polyploid white sturgeon (*Acipenser transmontamus*) using a dominant marker approach for duplicated microsatellite loci, Aquaculture 232 (2004) 165–182.
[35] P.A. Prodohl, W.J. Loughry, C.M. McDonough, W.S. Nelson, E.A. Thompson, J.C. Avise, Genetic maternity and paternity in a local population of armadillos assessed by microsatellite DNA markers and field data, Am. Nat. 151 (1998) 7–19.
[36] T.R. Meagher, E. Thompson, The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction, Theor. Popul. Biol. 29 (1986) 87–106.
[37] J.R. Gold, T.F. Turner, Population structure of red drum (*Sciaenops ocellatus*) in the northern Gulf of Mexico, as inferred from variation in nuclear-encoded microsatellites, Mar. Biol. 140 (2002) 249–265.
[38] J.M. Cornuet, S. Piry, G. Luikart, A. Estoup, M. Solignac, New methods employing multilocus genotypes to select or exclude populations as origins of individuals, Genetics 153 (1999) 1989–2000.