Genetic Structure and Diversity of the Endemic Carolina Madtom and Conservation Implications

W. Robert Cope
North Carolina Cooperative Fish and Wildlife Research Unit, Department of Applied Ecology, North Carolina State University, Campus Box 7617, Raleigh, North Carolina 27695, USA

Thomas J. Kwak*
U.S. Geological Survey, North Carolina Cooperative Fish and Wildlife Research Unit, Department of Applied Ecology, North Carolina State University, Campus Box 7617, Raleigh, North Carolina 27695, USA

Tyler R. Black
RK&K, 8601 Six Forks Road, Forum 1, Suite 700, Raleigh, North Carolina 27615, USA

Krishna Pacifici
Department of Forestry and Environmental Resources, North Carolina State University, 3120 Jordan Hall, Raleigh, North Carolina 27607, USA

Sheila C. Harris and Caitlin M. Miller
Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

Morgan E. Raley
HydroGENomics, Raleigh, North Carolina 27606, USA

Eric M. Hallerman
Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, USA

Abstract
Identification and conservation of genetic diversity within and among freshwater fish populations are important to better manage and conserve imperiled species. The Carolina Madtom Noturus furiosus is a small, nongame catfish that is endemic to the Tar and Neuse River basins of North Carolina. Genetic structure has not been studied in the species, and given recent population declines in both basins, identification of remaining genetic diversity within the species is vital for informing conservation efforts. To assess the status and trends of Carolina Madtom genetic
A recent study analyzed genetic markers from 173 individuals to (1) define population genetic structure, (2) assess intraspecific and interbasin genetic differentiation in the Tar and Neuse River basins, and (3) present management implications to guide conservation efforts. Using 10 microsatellite primers developed for the related Yellowfin Madtom *N. flavipinnis*, we observed low genetic diversity in Carolina Madtoms. Genotype frequencies within samples were not in Hardy–Weinberg equilibrium, with a deficit of heterozygotes that could be due to family structure, inbreeding, or segregation of null alleles. Mean ($\pm$SD) $M$-ratios for the Tar River (0.414 $\pm$ 0.117) and Neuse River (0.117 $\pm$ 0.102) basin collections indicated that both populations have experienced recent demographic bottlenecks, with that in the Neuse River basin population being more severe. Effective population size estimates for the respective populations were small, on the order of tens of individuals, driving low genetic diversity within populations. However, the multilocus population differentiation metrics $G_{ST}$ (mean $\pm$ SE = 0.135 $\pm$ 0.031) and $D_{EST}$ (0.125 $\pm$ 0.029) were significantly different from zero ($P < 0.001$), indicating significant genetic differentiation between the Tar and Neuse River basin populations. Our findings will inform managers on the status of genetic variation in the Carolina Madtom and will guide conservation toward protective listing and management decisions to maintain the viability of this important endemic species.

Identification and conservation of genetic diversity within and among freshwater fish populations are important to better manage and conserve imperiled species and preserve aquatic biodiversity. Imperiled species are especially vulnerable to loss of genetic diversity, often due to anthropogenic alterations to natural habitat (Vrijenhoek et al. 1985). Many factors, including urbanization, erosion, sedimentation, and pollution, degrade and destroy habitats that freshwater fishes require to survive and reproduce (Jelks et al. 2008). Dams cause population fragmentation and altered flow, both of which can negatively affect biodiversity and aquatic ecosystem health (Poff et al. 1997; Cooney and Kwak 2013). Agricultural impacts and urbanization can also fragment and degrade freshwater habitats. More than 40% of the area in six major river basins of the United States is considered agriculture land (Allan 2004). Urban land cover, while not composing large proportions of most river basins, can have disproportionately large effects on freshwater systems (Paul and Meyer 2001). Together, these impacts on altered landscapes can be detrimental to freshwater species, leading to decreased biodiversity and system health (Rahel 2002; Dudgeon et al. 2006; Kwak and Freeman 2010; Göthe et al. 2015).

Degraded and fragmented habitat directly affects fish by reducing the amount of available suitable habitat and divides species into multiple small populations. Fragmented populations may no longer intermix, and small, isolated populations are at risk of losing genetic diversity through inbreeding or random genetic drift, which can reduce the fitness and adaptability of such populations (Frankel and Soule 1981; Vrijenhoek et al. 1985). Identifying genetic structure and variation in such small populations is important for conservation efforts, especially with regard to imperiled and endemic species (Ryman 1991).

Loss of biodiversity and genetic variation is of special concern in the southeastern United States because many of the problems listed above, including damming, flow alteration, habitat degradation, and fragmentation, are common in the region (Jelks et al. 2008). Loss of genetic diversity is also of concern because the southeast region contains 62% of the freshwater fish species in the USA; however, it also contains the greatest number of imperiled fishes (Warren et al. 1997; Jelks et al. 2008). Many of these imperiled fishes are nongame species. Nongame fishes are critical to freshwater ecosystems by providing important ecological functions and contributing to healthy ecosystem biodiversity. However, many nongame fishes are in need of conservation action and are understudied and minimally managed compared to more economically or recreationally important species (Cooke et al. 2005; Kwak et al. 2011). As more nongame fishes face imperilment, the need for genetic conservation increases. In recent years, captive breeding and reintroduction efforts have become common among imperiled freshwater species, especially darters (*Ammocrypta*, *Crystallaria*, *Etheostoma*, and *Percina* spp.) and madtoms *Noturus* spp. (Rakes et al. 1999; Shute et al. 2005; Throneberry 2009). Thorough research on genetic structure among the extant populations should be conducted prior to conservation actions to assess genetic diversity among the few remaining populations. For example, reintroduction efforts for the Yellowfin Madtom *N. flavipinnis* and Smoky Madtom *N. baileyi* followed the development of microsatellite genetic markers and assessment of genetic differentiation among the parent and reintroduced populations (Williams and Moyer 2012). Such genetic markers also have been developed to quantify genetic diversity in many sport fishes as well as some other nongame fishes (Farias et al. 2003; Dutton et al. 2008; Hallerman et al. 2015). However, data on genetic diversity are lacking for many species, thus hampering conservation and management efforts.

One such understudied nongame species is the Carolina Madtom *N. furiosus*, a small catfish that is endemic to the Tar and Neuse River basins of North Carolina. Due to recent population declines, the Carolina Madtom is listed by the state of North Carolina as threatened...
GENETIC STRUCTURE OF THE CAROLINA MADTOM

(LeGrand et al. 2008; NCWRC 2014). However, the species was recently petitioned for federal protection by the U.S. Fish and Wildlife Service (USFWS) through the Endangered Species Act, and as such, federal and state protective listings may soon change (USFWS 2020). Although populations are declining and the species is becoming more imperiled, the Carolina Madtom has been sparsely studied, with only five major surveying events describing and assessing the status of extant populations (Bayless and Smith 1962; Smith and Bayless 1964; Burr et al. 1989; Wood and Nichols 2011; Cope 2018), and no population genetic research has been conducted. Neuse River basin populations in particular have been intensively impacted, with a 92% loss of historical occurrences in the basin over the past 50 years (Wood and Nichols 2011; Cope 2018). An extensive species status assessment conducted by USFWS personnel in 2018 corroborated these declines and evaluated the future viability of the Carolina Madtom, with most future scenarios leading toward basinwide extirpations (USFWS 2018). Thus, identification of the genetic diversity of the remaining populations in the two basins is critical for planning actions to conserve the species, as captive propagation has commenced and reintroduction or translocation efforts may be implemented.

Our research was designed to provide a timely assessment of the status and trends of Carolina Madtom genetic structure. Our objectives were to (1) define the population genetic structure of the Carolina Madtom, (2) assess intraspecific and interbasin genetic differentiation in the Tar and Neuse River basins, and (3) present management implications to guide conservation efforts for the species.

STUDY AREA

This research was conducted in the Tar and Neuse River basins of the Piedmont and Coastal Plain physiographic provinces in North Carolina (Figure 1). The Tar River runs through North Carolina from its origin in Person County to the town of Washington, where it becomes the Pamlico River and flows an additional 65 km to its mouth at Pamlico Sound (NCDWR 2001). The basin covers 8,755 km² and spans 16 counties (NCDWR 2001). With a human population of only 415,000, the Tar–Pamlico River basin is more rural and less impacted by human activities than the Neuse River basin. Currently, 55% of the Tar–Pamlico River basin is classified as forest and wetland, 28% is classified as crop and pasture, and 7% is classified as urban (NCDEQ 2010). The primary habitat problems affecting the basin are erosion and sedimentation due mainly to channel dredging for crop and livestock irrigation purposes (NCWRC 2005).

The Neuse River flows approximately 325 km through North Carolina from its headwaters originating in the Piedmont at the confluence of the Eno and Flat rivers to its mouth at Pamlico Sound near the city of New Bern (NCDWR 2010). The basin covers an area of 10,034 km² and spans 18 counties. Approximately 1.7 million people currently reside in the Neuse River basin, with the population expected to reach 3.0 million by the year 2050 (NCDWR 2010). Human activities in the river and its watershed impact the habitat and water quality of the Neuse River. Currently, 13% of the basin is considered urban, 45% is forested, and 29% is crop and pasture land (NCWRC 2005). In 2007, the American Rivers Foundation listed the Neuse River among the 10 most endangered rivers in the United States (American Rivers 2007). Non-point-source pollution from agriculture and forestry has degraded water quality and habitats throughout the basin. Commercial farming inputs, such as animal waste and fertilizers, contribute 60% of the nitrates and phosphates in the system (NCWRC 2005). Due to the dense human population in the basin, many municipalities have constructed dams and withdraw water from the impoundments for human use, thus affecting river flow, and treated sewage effluent is discharged into the river. Habitat loss is also an issue in the Neuse River basin, as the increasing human population results in a loss of natural areas and increases in impervious surfaces (NCDWR 2010).

METHODS

Collections.—Carolina Madtom populations in the Tar and Neuse River basins (Figure 1; Table 1) were sampled, and variation in their DNA was used to assess genetic diversity and population genetic structure. Genetic material used for analyses was obtained from four sources: tissue samples deposited in the North Carolina Museum of Natural Sciences (NCMNS; N = 6), Wood and Nichols (2011; N = 104), the North Carolina Wildlife Resources Commission (NCWRC; N = 21), and this study (N = 42). Samples from the NCMNS were collected during surveys in the early 2000s (2001–2007) and were stored in 95% ethanol in the museum’s fish collection. Samples from Wood and Nichols (2011) were collected during their basinwide Carolina Madtom surveys in summer 2007. The NCWRC samples were collected by NCWRC biologists in 2015–2017. Samples from the present study were collected in the summers of 2017 and 2018 during basinwide Carolina Madtom surveys at overlapping locations and by using a protocol similar to that of Wood and Nichols (2011). The field collection method for the majority of specimens was standardized snorkel surveys. During standardized snorkeling, 30–150-m transects were delineated at selected sites and were surveyed by snorkeling upstream along the transect and collecting all observed Carolina Madtom specimens by dip net (Wood and Nichols 2011; Cope 2018). Additional individuals were collected from
artificial cover units, which were constructed from terra cotta materials and deployed in river reaches at standardized snorkeling locations (Cope et al. 2019). Genetic samples were taken from Carolina Madtom individuals via a clip of the right pelvic fin before releasing the fish back into the water. Fin clips were placed into vials with 95% ethanol and were held at −20°C for storage until processing. All fin clips were deposited and documented in the NCMNS before being transported to Virginia Tech University, Blacksburg for molecular genetic analysis.

FIGURE 1. Sites in North Carolina where Carolina Madtom (CMT) collections were obtained for genetic analysis. Site numbers correspond to those described in Table 1. Specimens were collected from Fishing Creek (1–7), Little Fishing Creek (8), Sandy Creek (9), Swift Creek (10–13), Tar River (14–15), and Town Creek (16) in the Tar River basin and Contentnea Creek (17) and Little River (18) in the Neuse River basin.
Individuals from all collections were captured longitudinally along the river from upstream headwaters to the downstream confluence from all major tributaries in both basins (Figure 1). In the Tar River basin, collection sites included Fishing, Little Fishing, Swift, Town, and Sandy creeks and the main-stem Tar River. In the Neuse River basin, collection sites included Contentnea Creek and the Little River. Despite extensive surveying throughout the Neuse River basin, Carolina Madtoms were only found at the two sites included in our tissue sampling for this study (Contentnea Creek and Little River; Cope 2018).

**DNA markers.**—Extraction of DNA was performed using the Qiagen Blood and Tissue DNA Extraction Kit (Qiagen, Germantown, Maryland). The DNA concentration and purity were assessed using a μLite spectrophotometer (BioDrop, Cambridge, UK). The DNA was extracted from 173 individuals: 147 from the Tar River basin and 26 from the Neuse River basin. The panel of DNA samples was screened for variation at 11 microsatellite loci developed by Williams and Moyer (2012) for the congeneric Yellowfin Madtom. Amplification protocols followed those of Williams and Moyer (2012), modified as appropriate to yield optimized results using MyCycler thermocyclers (BioRad, Hercules, California). The 15-μL reactions consisted of 50 ng/μL of template DNA, 5× Promega reaction buffer, 2.00-mM MgCl₂, 1.25-mM dNTP mix, 0.5 μM of each forward and reverse primer, and 0.08 units of Promega Taq polymerase (Promega Corp., Madison, Wisconsin). Polymerase chain reactions followed the thermal profile of initial denaturation at 94°C (10 min), followed by a touchdown PCR protocol with 35 cycles of denaturation (94°C, 30 s), annealing, and extension (74°C, 30 s); the initial annealing temperature was 60°C (1 min) and decreased by 0.2°C per cycle. Final extension was at 74°C for 5 min. The presence of amplification products was checked by running an aliquot of reaction products through an ethidium bromide-stained tris–boric acid–EDTA agarose gel and observation of the gel under ultraviolet light. Multiplexed PCR products were subjected to fragment size analysis at the Cornell University Core Facility (Ithaca, New York) using an ABI 3730xl automated DNA sequencer (Applied Biosystems, Inc., Foster City, California).

**Data analysis.**—MICRO-CHECKER (van Oosterhout et al. 2004) was used to check the data for evidence of null alleles, large-allele dropout, or stuttering.

We used ML-Relate (Kalinowski et al. 2006) to perform maximum likelihood-based estimations of relatedness among individuals within the respective basins. Because Carolina Madtoms likely mate with individuals within the same subbasin, we repeated the analysis for collections with more than 20 individuals (i.e., for Swift Creek, Fishing Creek, the Tar River main stem, and Contentnea Creek).
We used program STRUCTURE version 2.3.4 (Pritchard et al. 2000), which utilizes a model-based approach to identify multilocus genotypic clusters within the data and to assign individuals to such multilocus clusters; this analysis is useful for informing the approach to subsequent statistical analyses. We utilized the admixture ancestry Bayesian model with correlated allele frequencies (Falush et al. 2003). We performed a burn-in of 10,000 runs and conducted 100,000 Markov chain Monte Carlo repetitions for \( K \) = 1–15 clusters, with five independent runs for each value of \( K \). The true number of populations (\( K \)) was assessed in two different ways. First, we sought the most likely \( K \) as indicated in the STRUCTURE output as the highest value of \( \log_{10}(P[X|K]) \)—that is, the probability (\( P \)) of observing the data (\( X \)) for a given level of \( K \). Second, we assessed the most likely \( K \) by calculating \( \Delta K \) using the Evanno et al. (2005) method. The method compares each \( K \) to \( \Delta K \), where \( \Delta K \) is calculated as mean\([L′(K)]/SD[L(K)]\]. The STRUCTURE analysis was performed twice: first using the whole data set, and second after removal of data for multiple members of full-sibling families and one instance of parent–offspring relationship and for samples in which fewer than 20 individuals were collected. In an attempt at visualizing genetic clustering by a method that was not dependent upon Hardy–Weinberg equilibrium, we applied discriminant analysis of principal components (Jombart et al. 2010) using adegenet version 2.1.3 (Jombart et al. 2020).

Polymorphism in individual populations and for data pooled across populations was quantified as the number of alleles per locus (\( A \)), observed heterozygosity (\( H_O \)), and expected heterozygosity (\( H_E \)) using the Microsatellite Toolkit (Park 2001) and Arlequin version 3.5.1.3 (Excoffier and Lischer 2010). Arlequin was used to assess linkage disequilibrium (LD) with 10,000 permutations. Arlequin was used to apply the Hardy–Weinberg exact test (Haldane 1954) to the null hypothesis of random union of gametes. A Markov chain method with 10,000 dememorization steps and 100,172 iterations was used to estimate, without bias, the exact \( P \)-value for this test (Levene 1949; Guo and Thompson 1992). All \( P \)-values were adjusted for multiple comparisons using the sequential Bonferroni method (Rice 1989).

Analysis of molecular variance (AMOVA) and calculation of \( F \)-statistics followed the methods of Weir and Cockerham (1984), Excoffier et al. (1992), and Weir (1996) and were performed using Arlequin with 10,000 permutations. Levels of hierarchy within the AMOVA were as follows: within individuals, among individuals within populations (Tar or Neuse River basin), and among populations (Tar and Neuse River basins). Calculation of \( F \)-statistics and \( R_{ST} \) (a statistic for measuring genetic distance between subgroups or populations using the stepwise mutation process of microsatellites; Slatkin 1995) was performed in Arlequin.

Using GenAIEx version 6.5 (Peakall and Smouse 2012), we calculated \( G_{ST} \) (a standardized measure of genetic differentiation, which has a range from 0 to 1 for all loci, independent of the extent of within-population genetic variation; Hedrick 2005) and \( D_{ST} \) (a standardized measure of genetic differentiation, which has a range from 0 to 1 for all loci, independent of the extent of within-subpopulation genetic variation; Jost 2008), and their significance was assessed using 1,000 permutations of the data. Significance tests were performed with 10,000 permutations. The Garza–Williamson index (\( M \)-ratio) is a microsatellite marker-based indicator of recent genetic population bottlenecks in a population (Garza and Williamson 2001). The \( M \)-ratio for each locus—that is, the number of alleles exhibited by a population at a particular locus as a proportion of alleles possible within the allelic size range at that locus (Garza and Williamson 2001)—was calculated using Arlequin. The index indicates a recent genetic bottleneck at values below 0.70 when sample size is less than 75.

Effective population size (\( N_e \)) and significant LD were estimated using LDNe version 1.31 (with the random mating model, lowest allele frequency of 0.02, and parametric CIs; Waples and Do 2008) and NeEstimator (Do et al. 2014). These software applications utilize the bias-corrected LD (Waples and Do 2008), molecular coancestry (Nomura 2008), and heterozygosity excess (Zhdanova and Pudovkin 2008) methods.

**RESULTS**

All 11 microsatellite primer pairs developed for the Yellowfin Madtom amplified Carolina Madtom DNA. Amplification products of 10 loci were readily interpretable (Table SI 1 in Cope 2018); however, results for locus \( Nf/D139 \) were not interpretable.

Analyzed globally using MICRO-CHECKER, genotypic frequencies at the 10 loci scored were not in Hardy–Weinberg equilibrium (Table 2), suggesting segregation of null alleles or the effects of population genetic processes, such as inbreeding or a Wahlund effect (i.e., the mixing of differentiated populations). Significant LD was observed at 79 of 81 locus-by-locus comparisons in the Tar River population and 65 of 81 comparisons in the Neuse River population. It is highly unlikely that all of the microsatellite loci screened are linked in the sense of being located on the same chromosome. Linkage disequilibrium can result from several processes operating singly or in combination, including epistatic selection, migration, genetic hitchhiking, or random drift in finite populations (Hedrick et al. 1978; Hill 1981). Because the individuals that we genotyped were drawn from small populations, genotypes at independent loci may have shown apparent linkage because sampled individuals descended from a limited number of breeders.
of Locus = observed heterozygosity; populations of Carolina Madtoms (Table 3) was less than 10,440 pairwise comparisons, there were 1 parent–offspring, 12 full-sibling, and 14 half-sibling pairs; 7.1% of comparisons were among related individuals. Conducting the analysis at the subbasin level, we found considerable family structure (Table 4), especially for collections in the Tar River basin. For example, among 52 individuals within the Fishing Creek collection, we identified 12 full-sibling families with 2–10 members each; half-siblings of these individuals added family members. The family structure within the data likely contributed to departure from Hardy–Weinberg equilibrium and to positive $F_{IS}$ values (the departure from Hardy–Weinberg expected genotype frequencies due to within-population processes).

Results of STRUCTURE analysis for the full data set showed that the log$_{10}$(Pr[X|K]) criterion declined as $K$ increased well beyond the number of collections made (Cope 2018). The Evanno et al. (2005) $\Delta K$ criterion indicated that a $K$-value of 2 was the best-supported number of clusters. At $K=2$, there was no geographic basis for distribution of genetic variation; at higher levels of $K$, differentiation among individuals—but not geographically defined collections—became more narrowly drawn. This pattern of results suggests a lack of geographically based population-level genetic differentiation in the Carolina Madtom. Because our STRUCTURE results may have been affected by occurrence of null alleles, inclusion of multiple members of families in the data set (Anderson and Dunham 2008; Rodriguez-Ramilo and Wang 2012), or uneven sample sizes (Gilbert 2016; Janes et al. 2016; Puechmaille 2016), we reanalyzed our data using a reduced data set that included only one member of each full-sibling family and collections with sample sizes of 20 or more. The results for the reduced data set were similar to those for the full data set inasmuch as there was no clear geographic basis for clustering. Using the Evanno et al. (2005) criterion, the best support was for a $K$-value of 2 clusters; for the log$_{10}$(Pr[X|K]) criterion, a $K$-value of 9 received the most support.

The results from application of the discriminant analysis of principal components showed the greatest support for five clusters of individuals based on the Bayesian information criterion. The clustering of individuals, however, had no particular geographic basis.

Results of AMOVA (Table 5) showed that most of the genetic variation (98.04%) was partitioned within the respective populations and rather little (1.96%) was partitioned between the Tar and Neuse River basin populations. All $F$-statistics were significantly different from zero. A metric of total departure of genotype frequencies from Hardy–Weinberg equilibrium expectations ($F_{IT}$) was 0.502, a reasonably high value. The index $F_{IS}$, which quantifies within-population departures of genotype

Table 2. Results of MICRO-CHECKER analysis of genotype frequencies at 10 microsatellite loci in Carolina Madtoms, with estimated frequency of null alleles obtained using the van Oosterhout et al. (2004) method.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Null alleles present</th>
<th>Estimated frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>NflA3</td>
<td>Yes</td>
<td>0.208</td>
</tr>
<tr>
<td>NflC138</td>
<td>Yes</td>
<td>0.250</td>
</tr>
<tr>
<td>NflC143</td>
<td>Yes</td>
<td>0.473</td>
</tr>
<tr>
<td>NflC145</td>
<td>Yes</td>
<td>0.472</td>
</tr>
<tr>
<td>NflD105</td>
<td>Yes</td>
<td>0.380</td>
</tr>
<tr>
<td>NflD123</td>
<td>Yes</td>
<td>0.095</td>
</tr>
<tr>
<td>NflD129</td>
<td>Yes</td>
<td>0.243</td>
</tr>
<tr>
<td>NflD137</td>
<td>Yes</td>
<td>0.263</td>
</tr>
<tr>
<td>NflD145</td>
<td>Yes</td>
<td>0.120</td>
</tr>
<tr>
<td>NflD146</td>
<td>Yes</td>
<td>0.198</td>
</tr>
</tbody>
</table>

All microsatellite loci were variable for Carolina Madtoms in each drainage (Table 3), with a mean ($\pm$SD) $A$ of 17.7 $\pm$ 7.7 in the Tar River basin and 11.8 $\pm$ 7.3 in the Neuse River basin. For all loci in both populations, $H_O$ was less than $H_E$, indicating significant ($P < 0.001$) departures from Hardy–Weinberg equilibrium expectations. This finding may be attributable to segregation of null alleles, family structure within the data, inbreeding, or any combination of these processes.

Maximum likelihood-based analysis showed considerable frequencies of related individuals within each of the respective basins. Within the Tar River basin, among 10,440 pairwise comparisons, there were 0 parent–offspring, 674 full-sibling, and 649 half-sibling pairs; thus, 12.6% of paired comparisons were among related individuals. Within the Neuse River basin, among 378 pairwise comparisons, there were 1 parent–offspring, 12 full-sibling, and 14 half-sibling pairs; 7.1% of comparisons were among related individuals. Conducting the analysis at the subbasin level, we found considerable family structure (Table 4), especially for collections in the Tar River basin. For example, among 52 individuals within the Fishing Creek collection, we identified 12 full-sibling families with 2–10 members each; half-siblings of these individuals added family members. The family structure within the data likely contributed to departure from Hardy–Weinberg equilibrium and to positive $F_{IS}$ values (the departure from Hardy–Weinberg expected genotype frequencies due to within-population processes).

Table 3. Metrics of genetic variation in Tar and Neuse River basin populations of Carolina Madtoms ($A = $ number of alleles per locus; $H_O = $ observed heterozygosity; $H_E = $ expected heterozygosity). All departures of $H_O$ from $H_E$ were statistically significant ($P < 0.001$).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Tar River basin</th>
<th>Neuse River basin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A$</td>
<td>$H_O$</td>
</tr>
<tr>
<td>NflA3</td>
<td>12</td>
<td>0.485</td>
</tr>
<tr>
<td>NflC138</td>
<td>10</td>
<td>0.023</td>
</tr>
<tr>
<td>NflC143</td>
<td>8</td>
<td>0.000</td>
</tr>
<tr>
<td>NflD105</td>
<td>21</td>
<td>0.200</td>
</tr>
<tr>
<td>NflD123</td>
<td>34</td>
<td>0.771</td>
</tr>
<tr>
<td>NflD129</td>
<td>15</td>
<td>0.458</td>
</tr>
<tr>
<td>NflD137</td>
<td>13</td>
<td>0.390</td>
</tr>
<tr>
<td>NflD138</td>
<td>14</td>
<td>0.345</td>
</tr>
<tr>
<td>NflD145</td>
<td>23</td>
<td>0.746</td>
</tr>
<tr>
<td>NflD146</td>
<td>20</td>
<td>0.561</td>
</tr>
<tr>
<td>Mean</td>
<td>17.7</td>
<td>0.409</td>
</tr>
<tr>
<td>SD</td>
<td>7.7</td>
<td>0.266</td>
</tr>
</tbody>
</table>
frequencies from Hardy–Weinberg expectations, was 0.490; the fact that \( F_{IS} \) was positive indicated a rather large deficit of heterozygotes. The value of \( F_{IS} \) was somewhat greater in the Neuse River population (0.520) than in the Tar River population (0.486) of Carolina Madtoms. Using this algorithm, a metric of population-level differentiation (\( F_{ST} \)) was 0.019, which is rather low. A metric of population-level differentiation that assumes stepwise mutation of microsatellite alleles and includes consideration of allele size (\( R_{ST} \)) was somewhat higher than \( F_{ST} \), as expected (\( R_{ST} = 0.035 \)), but remained rather low. Using metrics developed for quantifying population-level differentiation based on microsatellite loci, \( G_{ST} \) was 0.135 ± 0.031 (mean ± SE) and \( D_{EST} \) was 0.125 ± 0.029; both were higher than \( F_{ST} \), as expected, and both were significantly different from zero (\( P < 0.001 \)).

We repeated the analysis using four subpopulations with samples of at least 20 individuals (i.e., with separate consideration of Swift Creek, Fishing Creek, and the Tar River main stem [Tar River basin] and Contentnea Creek [Neuse River basin]). Overall \( F_{ST} \) was 0.032 ± 0.004 (mean ± SE), \( G_{ST}^\prime \) was 0.137 ± 0.031, and \( D_{EST} \) was 0.125 ± 0.029. Matrices of subpopulation-by-subpopulation \( G_{ST}^\prime \) and \( D_{EST} \) values (Table 6) showed that many such pairwise metrics of genetic differentiation were significantly different from zero.

In the original work of Garza and Williamson (2001), \( M \)-ratios averaged 0.87 for demographically stable populations and 0.64 for bottlenecked populations. Low \( M \)-ratios (Table 7) for Carolina Madtoms in the Tar River (mean ± SD = 0.414 ± 0.117) and Neuse River (0.341 ± 0.102) basins indicated that both populations have experienced recent demographic bottlenecks, with the Neuse River basin population experiencing a more severe one.

Using the LD method, \( N_e \) estimates for the respective populations were small—on the order of tens of individuals (Table 8). Molecular coancestry-derived estimates of the effective number of breeders (\( N_{eb} \)) were 11.0 (jackknifed 95% CI = 5.6–18.2) for the overall system, 10.5 (5.4–17.1) for the Tar River drainage, and 12.7 (3.5–27.8) for the Neuse River drainage. Use of the heterozygosity excess method led to unresolved \( N_e \) estimates, with CIs including zero and infinity.

### DISCUSSION

Carolina Madtom genetic variation was observed using 10 microsatellite markers from the congeneric Yellowfin Madtom. Although cross-species primer transfer was shown to be relatively unsuccessful for the related Smoky Madtom, for which only three primer pairs successfully amplified, we successfully amplified and attained readable results from 10 of 11 tested primer pairs. Cross-species transferability of microsatellite primers is highly variable, and successful amplification is not common; as such, these results are promising for interspecies analyses (Scribner et al. 1996; Barbara et al. 2007). However, as is the case in most cross-species primer transfers, genetic results were

---

**TABLE 4.** Maximum likelihood-estimated numbers of inferred relationships among Carolina Madtoms within collections of 20 or more individuals.

<table>
<thead>
<tr>
<th>Waterway</th>
<th>N</th>
<th>Comparisons</th>
<th>Parent–offspring</th>
<th>Full-sibling</th>
<th>Half-sibling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tar River basin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swift Creek</td>
<td>56</td>
<td>1,301</td>
<td>0</td>
<td>47</td>
<td>83</td>
</tr>
<tr>
<td>Fishing Creek</td>
<td>52</td>
<td>1,177</td>
<td>0</td>
<td>60</td>
<td>89</td>
</tr>
<tr>
<td>Tar River main stem</td>
<td>33</td>
<td>478</td>
<td>1</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>Neuse River basin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contentnea Creek</td>
<td>20</td>
<td>219</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

**TABLE 5.** Analysis of molecular variance results quantifying percentages of genetic variation at hierarchical levels within Carolina Madtom individuals, populations, and overall. Populations are defined as the Tar and Neuse River basins.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td>1</td>
<td>7.164</td>
<td>0.04248</td>
<td>1.96</td>
</tr>
<tr>
<td>Among individuals within populations</td>
<td>171</td>
<td>543.159</td>
<td>1.04772</td>
<td>48.26</td>
</tr>
<tr>
<td>Within individuals</td>
<td>173</td>
<td>187.00</td>
<td>1.08092</td>
<td>49.79</td>
</tr>
<tr>
<td>Total</td>
<td>345</td>
<td>737.324</td>
<td>2.17113</td>
<td></td>
</tr>
</tbody>
</table>

---
not perfect. Genotype frequencies from our Carolina Madtom samples were not in Hardy–Weinberg equilibrium, with a severe lack of heterozygous genotypes. At a population genetic level, this outcome could be explained by family structure or inbreeding, which increases the frequency of homozygotes and reduces the frequency of heterozygotes within a population. However, we cannot unequivocally conclude that our results are due to these population genetic processes, as the cross-species primer pairs returned high frequencies of null alleles, which could create the false appearance of excess homozygotes in the population. Family structure within our collections also could have contributed to departure from Hardy–Weinberg equilibrium.

Using these genetic markers, we found that Carolina Madtoms had low genetic diversity throughout both the Tar and Neuse River basins, as determined through multiple testing methods. Carolina Madtoms are common in isolated locations in the Tar River basin and would be expected to show moderate levels of genetic variability. However, populations in the Neuse River basin have declined demographically, and the screened samples would be expected to harbor less genetic variability. This outcome was apparent, with populations from the Neuse River basin displaying reduced numbers of alleles and lower \( M \)-ratios than Tar River basin populations. The \( M \)-ratios for populations in both basins indicated recent genetic bottlenecks, with the Neuse River basin experiencing a slightly more severe bottleneck. Furthermore, collections in both basins, especially the Tar River basin, showed family structure, suggesting limited numbers of breeders. Both basins also exhibited low values of \( N_e \) on the order of tens of individuals, explaining the low levels of genetic diversity within populations. However, using measures of genetic variation among groupings of collections within subbasins, we found considerable genetic differentiation among the Tar River, Fishing Creek, Swift Creek, and Contentnea Creek populations, suggesting strong effects of random genetic drift.

Results of maximum likelihood-based relatedness analysis showed reasonably high frequencies of relationships among individuals within the respective basins. These results support the interpretation that the populations experienced demographic bottlenecks and that inbreeding is ongoing.

Our study is the first to describe the genetic structure of the Carolina Madtom. Identification of genetic structure and variability is critical for management and conservation decisions for a species. Identification of such genetic variability is especially important among endangered and endemic species, which are at greatest risk of a sudden loss of populations and genetic variability (Ryman 1991).
Assessment of genetic variability is commonly practiced to define management plans for imperiled species, and the results for the Carolina Madtom are not unique. Identification of genetic variation in Shortnose Sturgeon Acipenser brevirostrum, Yazoo Darter E. raneyi, and the Neotropical catfish Steindachneridion parahybae have shown that anthropogenic habitat alteration has negatively affected the genetic diversity of the species (Quattro et al. 2002; Sterling et al. 2012; da Fonseca et al. 2017). Research on the Caddo Madtom Noturus taylori has also shown similarly low genetic diversity due to small, fragmented populations (Turner and Robison 2006). However, much like the Carolina Madtom, Caddo Madtoms exhibited structuring among individual subpopulations, even at a relatively small spatial scale. Benthic habitat specialist species, such as the Brindled Madtom N. miurus and Paleback Darter E. pallididorsum, are likely to exhibit genetic variation at a relatively small spatial scale because of their inability to migrate between populations separated by patches of unsuitable habitat. Such genetic variation was seen in the Carolina Madtom, as three collections from separate subbasins in the Tar River basin exhibited significant genetic differentiation, although they are not greatly separated spatially (Fishing and Swift creeks are within 11 km of each other off the main-stem Tar River). Interbasin genetic variation is also common in imperiled fishes. The Boga-portuguesa Chondrostoma lusitanicum, an imperiled, endemic cyprinid, exhibited significant genetic variation across basins (Mesquita et al. 2001). Diadromous species, such as smelt Retropinna spp., also have exhibited interbasin genetic variability, even though their marine larval stage allows for connectivity among multiple river basins (Hughes et al. 2014). Connectivity between the Tar and Neuse rivers likely last occurred during the Pleistocene epoch, when the sea level was 100 m lower than it is currently (Bloom 1983). With the rise of sea level, Carolina Madtom populations in the two river basins became demographically independent and subject to independent evolutionary histories.

Conservation Implications

Genomic method.—Although we successfully amplified primers developed for the congeneric Yellowfin Madtom, development and use of species-specific microsatellite primers for the Carolina Madtom would allow for greater confidence in genetic analyses and provide greater accuracy and genetic resolution to guide conservation of the species. The segregation of null alleles was apparent in our data set, which impacted our ability to reach unequivocal inferences of population genetic structure and underlying processes in the Carolina Madtom. As such, microsatellite locus primer pairs developed specifically for Carolina Madtoms would allow for more accurate analyses of microsatellite DNA data.

Application of genomic methods—that is, screening of a large collection of single-nucleotide polymorphisms—could yield new insights useful for planning and implementing conservation actions for the Carolina Madtom. Screening of single-nucleotide polymorphism haplotypes can support robust assessment of inbreeding (Allendorf et al. 2010; Leroy et al. 2018), LD-based estimation of N_e (Waples et al. 2016), demonstration of population genetic differentiation (Larson et al. 2014), identification of adaptive genetic variation (Allendorf et al. 2010; Hohenlohe et al. 2010; Moore et al. 2014), and delineation of conservation units (Funk et al. 2012; McKinney et al. 2020). Analysis of such data could clarify the causality of the departures from Hardy–Weinberg equilibrium that we observed and could provide insights into historical demography and, in turn, the degree to which population genetic differentiation is longstanding and natural or the result of recent bottlenecks and random genetic drift. New tools for analyzing genomic data are useful for assessing population genetic processes that are ongoing in declining populations (Leroy et al. 2018).

Species conservation.—Given the results from our study, a variety of management options may be available to conserve the remaining genetic variation of the Carolina Madtom. The Carolina Madtom fits the classical, small-population paradigm of conservation genetics (Caughley 1994): its populations are isolated and small and, hence, are subject to the effects of random drift-mediated loss of genetic variation and inbreeding. Genetic differentiation, such as that indicated by the significant G_ST and D_FST metrics that we observed, may be due to random processes as opposed to adaptive processes.

Captive breeding and propagation of Carolina Madtoms have been initiated on a limited scale to produce fish for possible use in reintroduction into historical habitat where they have been extirpated or in augmentation of existing populations. Reintroduction efforts are common among federally listed threatened and endangered species as viable population recovery goals (Tear et al. 1993). However, concerns have been raised over reintroduction efforts, as species often are reintroduced into habitat that retains the problematic factors that caused the original population collapse (Conant 1988). Even with such concerns, multiple madtom species have been successfully reintroduced in the southeastern United States. The Smoky and Yellowfin madtoms have been successfully propagated in captivity and released into stream habitat in their native ranges in Tennessee; since their reintroduction, successful reproduction and population increases in the wild have been documented (Rakes et al. 1999; Shute et al. 2005; Throneberry 2009). Given these findings,
reintroduction may prove a practical management alternative for Carolina Madtoms.

Conservation planning for reintroduction or augmentation activities involves consideration of balancing the extinction risks from inaction with the demographic or genetic benefits and the genetic risks of potential actions (Miller and Kapuscinski 2003; George et al. 2009). In the case of translocations into or augmentation of an existing population, the benefits of genetic rescue (Frankham 2015; Bell et al. 2019) and the risk of outbreeding depression (Frankham et al. 2011) must be weighed. Although restoring gene flow into small, isolated populations can reduce the impact of inbreeding and decrease extinction risk (i.e., effect genetic rescue; Bell et al. 2019), human-mediated gene flow has only rarely been utilized as a strategy for conservation management of fishes (Fitzpatrick et al. 2016; Robinson et al. 2017). Our results for the Carolina Madtom indicate the impact of a drift-mediated loss of genetic variation, and our observations of limited numbers of families at given sites and high $F_{IS}$ values suggest the potential for inbreeding. Populations within the Tar River basin likely have become isolated only in historic times, suggesting only a modest risk of outbreeding depression if distinct populations are brought into genetic contact by translocation or augmentation (Frankham et al. 2011). The Neuse and Tar River basin populations have long been separated, and while the relative risks of inbreeding and outbreeding depression might be weighted differently than for within-basin transfers, loss of genetic variation and inbreeding are genetic risks that might be addressed by well-designed conservation interventions. Translocation or stocking of captive-propagated fish to establish “experimental” populations, in the sense defined in the federal Endangered Species Act, might be suitable for assessing the utility of mixing populations for purposes of conservation management.

Prior to any population reintroduction or augmentation activities in the impacted Neuse River basin, identification and amelioration of the stressors and impacts that caused the population collapse in the basin would be required. Human-mediated impacts to the Neuse River basin are impediments to conservation of the Carolina Madtom. Despite extensive sampling efforts, the subbasin population from the Neuse River basin at Contentnea Creek had not been detected in over 10 years until recent sampling efforts by NCWRC personnel in 2018 yielded one individual. However, no genetic material was collected from the individual, leaving all known genetic diversity to come from a small number of individuals occurring in the Little River. The threats facing the Neuse River basin also make reintroduction or augmentation efforts a challenge. Non-native Flathead Catfish *Pylodictis olivaris* have been introduced into the basin (Kwak et al. 2006). Flathead Catfish are voracious predators that have been shown to suppress native fish populations by up to 50% and to positively select ictalurid prey, including native bullheads *Ameiurus* spp. and madtom species in North Carolina rivers (Pine et al. 2005, 2007; Baumann and Kwak 2011). Snorkel surveying during our study documented co-occurrence of Carolina Madtoms and Flathead Catfish in the Little River, which may be harboring the last individuals in the entire Neuse River basin (Cope 2018; Cope et al. 2019). Therefore, reintroduction of captively propagated Carolina Madtoms may not be advisable because any reintroduced individuals would be at risk of immediate predation by Flathead Catfish unless released into experimental locations where Flathead Catfish are not known to be abundant or to occur.

With the threat of Flathead Catfish predation in the Neuse River basin, a reasonable approach to population expansion may be the establishment of experimental populations from captively propagated individuals in reaches away from areas where predation is likely. Timely collection of the remaining spawning adults in the Little River may be considered to provide broodstock for reintroduction purposes. Using captively propagated individuals from the Neuse River basin may allow for the introduction of Carolina Madtoms into areas with suitable habitat that are currently not occupied by Flathead Catfish to determine whether captively propagated individuals can successfully survive and reproduce in the wild. Experimental populations may also be informative in testing the survival, viability, and adaptability of various cross-bred progenies between individuals from multiple basins and subbasins. Historical Carolina Madtom occurrences have been recorded in the Eno River of the upper Neuse River basin above Falls Lake, north of Raleigh, North Carolina (Wood and Nichols 2011). This location, as well as the adjoining Flat River, may be suitable for an experimental Neuse River basin population, as there is recorded suitable habitat for Carolina Madtom occurrence and these tributaries lie upstream of major municipal effluents and other impacts. Human population growth and increased anthropogenic land use, such as those in the Tar and Neuse River basins, can lead to increased runoff from artificial impervious surfaces and a greater number of point and nonpoint sources of pollutants entering the water system (e.g., sewage, industrial effluent, and road runoff); metals and other aquatic contaminants have been identified as potential factors leading to population decreases in other madtom species (Wildhaber et al. 2000; Allen et al. 2001; Cope 2018). Additionally, the creation and damming of the Falls Lake reservoir (area $= 5,022 \text{ ha}$) have prevented upstream movement of the abundant Flathead Catfish population from the downstream Neuse River into the reservoir and upstream Eno and Flat rivers. As of summer 2020, Flathead Catfish have not been documented in Falls Lake, the Eno River, or the Flat River during routine sampling of those water bodies by NCWRC biologists (K. Roberts and
ACKNOWLEDGMENTS

Funding for this research was provided by the NCWRC through the State Wildlife Grant Program. We thank William Wood, Joseph McIver, Spencer Gardner, Tom Fox, Zoe Nichols, and Mike Walter for field and laboratory support. The manuscript was strengthened by attention to the comments of the anonymous peer reviewers. The participation of co-author Eric M. Hallerman was supported in part by the Virginia Agricultural Experiment Station under the U.S. Department of Agriculture. The participation of co-author Eric M. Hallerman also greatly benefited from the attention to the comments of the anonymous peer reviewers. William Wood, Joseph McIver, Spencer Gardner, and K. Rundle, NCWRC, personal communication. The reduction or lack of known major stressors to the Carolina Madtom in the Eno and Flat rivers indicates that these tributaries may be suitable locations for experimental population reintroductions.

The Carolina Madtom is an important component of the stream ecosystem, functioning both as an important ecological link and as part of the network of endemics that render North Carolina and southeastern U.S. freshwater systems biologically diverse and unique. Given the species’ recent decline and current small, fragmented populations, conservation is vital not only to maintain a unique, endemic genome, but also to maintain the diversity and ecological integrity of these waterways. The application of our results will inform managers on the status of genetic variation in Carolina Madtoms and will guide protective listing, planning, and decisions and actions to maintain the viability of this endemic species with respect to both demographic and genetic variation needs to increase targeted populations.

REFERENCES


Puechmaile, S. J. 2016. The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Molecular Ecology Resources 16:608–627.


