Genetic Monitoring of Wild and Captive Populations of Barrens Topminnow (Fundulus Julisia)

Article *in* Southeastern Naturalist · June 2025

DOI: 10.1656/058.024.0sp1301

4 authors, including:

Carla Hurt
Tennessee Technological University
80 PUBLICATIONS **1,266** CITATIONS

SEE PROFILE

Bernard Kuhajda
Tennessee Aquarium Conservation Institute
76 PUBLICATIONS **1,568** CITATIONS

SEE PROFILE

Genetic Monitoring of Wild and Captive Populations of Barrens Topminnow (*Fundulus julisia*)

Carla Hurt[1,*], Bernard Kuhajda[2], Lydia Burton[1,3], and Bennett Agee[1]

Abstract - Over the past 30 years, captive propagation and reintroduction efforts have been undertaken to mitigate the decline of the federally endangered *Fundulus julisia* (Barrens Topminnow [BTM]) throughout the Barrens Plateau region of Middle Tennessee. However, concerns remain about the loss of genetic diversity and adaptive potential in captivity, especially as reintroductions of captive-reared fish have had mixed success. This study analyzed genome-wide SNP genotypes in captive and wild BTM populations to assess patterns of genetic variation in captive populations relative to their wild counterparts. Estimates of heterozygosity and pairwise population distances indicated significant genetic drift and reduced genetic diversity when compared with source populations. Notably, a newly discovered natural population of BTM in the upper Collins River was found to be genetically distinct, warranting its classification as a separate evolutionarily significant unit. These findings can provide insights into the management of the BTM captive-breeding program and underscore the need for protection of the upper Collins River population.

Introduction

Populations of *Fundulus julisia* Williams and Etnier (Barrens Topminnow [BTM]) were once widespread throughout the headwaters of the Elk River, Duck River, and Collins River drainages in Middle Tennessee, but over the last 50 years, droughts, widespread introduction of the invasive *Gambusia affinis* (Baird and Girard) (Western Mosquitofish), and human disturbances have resulted in their dramatic decline (Laha and Mattingly 2007, USFWS 2024a). Captive populations of BTM were initiated by brood stock from native wild (natural) populations in the Elk River and Collins River (Caney Fork) drainages and continue to be maintained at Conservation Fisheries Inc. (CFI), the Tennessee Aquarium (TNACI), the Dale Hollow Fish Hatchery (DHFH), and the Wolf Creek Fish Hatchery (WCFH). These captive-reared fish are currently being used as sources for reintroductions at selected sites within the former range of the species. However, captive populations have been genetically isolated from natural populations for multiple generations, and little is known about the current genetic composition of captive populations relative to their natural source populations.

Natural populations of BTM have been reduced to 4 extant sites, including a recently discovered population in the upper Collins River (Fig. 1; Kuhajda et al. 2014; W. Stiles, USFWS, Cookeville, TN, pers. comm.). These natural populations,

[1]Department of Biology, Tennessee Technological University, Cookeville TN 38505. [2]Tennessee Aquarium Conservation Institute, Chattanooga, TN 37405. [3]The Department of Applied Environmental Science, California State University, Monterey Bay, CA 93955. [*]Corresponding author - churt@tntech.edu.

Manuscript Editor: David Duvernell

exclusive of the upper Collins River population, have been used to establish 3 different captive-bred lineages. Populations at Pedigo Farm and Pedigo Highway in the Lewis/McMahan creek system in the Collins River watershed were the source for brood stock for the Lewis Creek lineage. The Benedict Spring population (type locality) in the Hickory Creek system (also in the Collins River watershed) served as the source for brood stock for the second lineage; this population is believed to

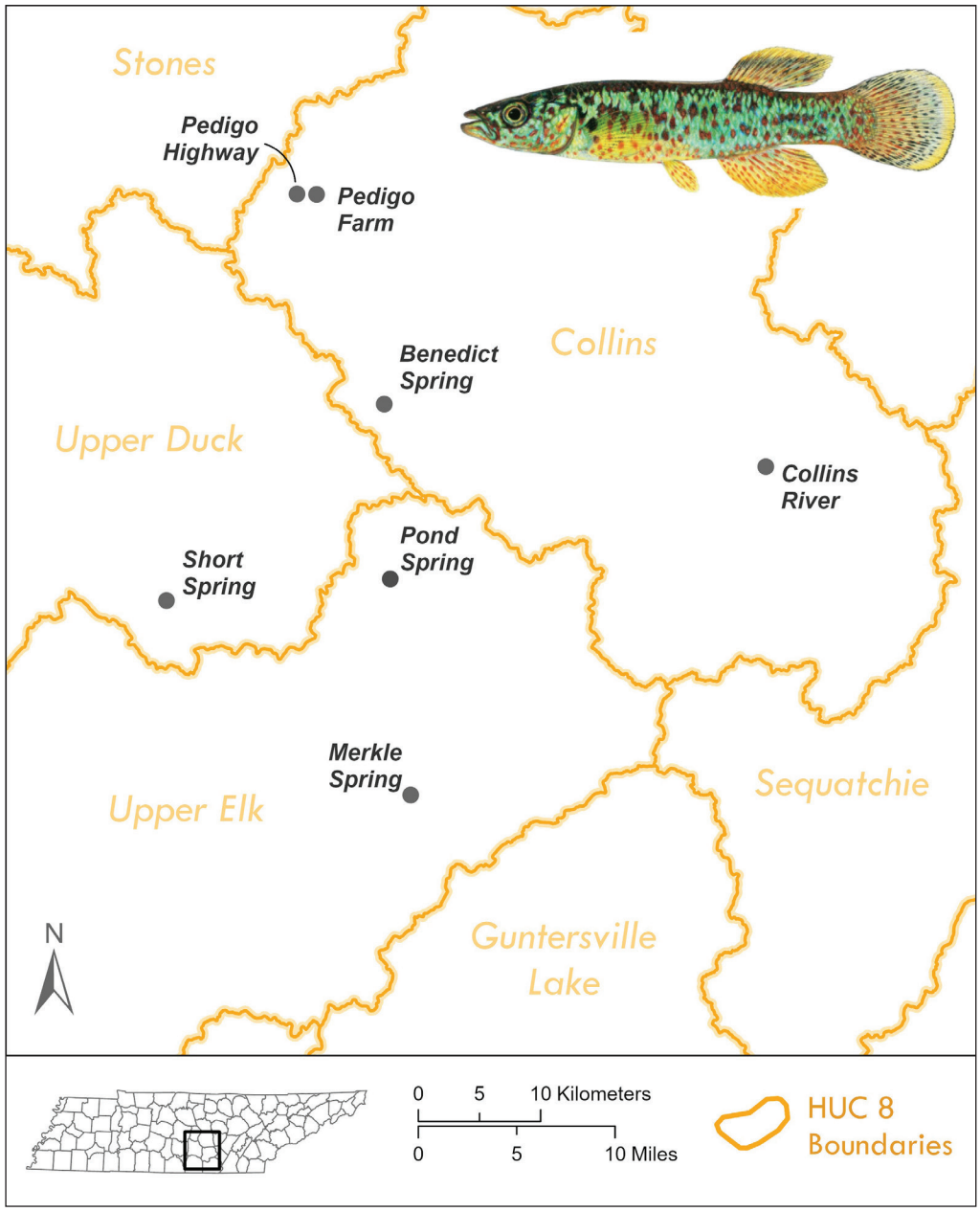

Figure 1. Map of Barrens Topminnow populations represented in this study. Inset shows an illustration of a male Barrens Topminnow in breeding colors. Illustration © Joe Tomelleri.

have been extirpated in 2022 after a repeated history of droughts (Hurt et al. 2017). Lastly, a population at Pond Spring, thought to have been extirpated between 2012 and 2013 (Kuhajda and Mitchell 2019), provided brood stock for the Elk River lineage. Pond Spring was the only known population of BTM to occur in the Elk River watershed. A survey of genetic variation at 14 microsatellite loci and mitochondrial sequence data demonstrated that these 3 lineages are genetically distinct (Hurt et al. 2017). This study identified Pond Spring as its own evolutionary significant unit (ESU), where ESUs are defined as historically isolated groups of populations that are on independent evolutionary trajectories as indicated by both mitochondrial and nuclear loci (Moritz 1994). Populations in the Collins River/Caney Fork watershed represented a second ESU; Pedigo/Lewis (Witty Creek) and Benedict Spring/Type (Hickory Creek) were assigned to separate management units (MUs), which are less distinct than ESUs and only need to demonstrate significant divergence in allele frequencies at nuclear loci. These 3 lineages have been maintained separately at CFI, TNACI, DHFH, and WCFH and have been used for reintroductions throughout the BTM's historical range.

Conservation efforts for BTMs started as early as the 1970s (USFWS 2024c), but coordinated propagation and stocking programs did not begin until 2001 with the formation of the Barrens Topminnow Working Group. This working group is represented by state and federal agencies, universities, and non-profit organizations. At this time, natural brood stock were available from native populations in the Elk River drainage at Pond Spring and at several sites in the Collins River drainage, including the Pedigo Farm (Lewis Creek) and Woodland Estate sites in the Witty Creek MU and Benedict Spring (type locality) in the Hickory Creek MU. CFI and TNACI propagated BTMs, and offspring were grown out at these 2 facilities as well as DHFH and WCFH (Figs. 1, 2; USFWS 2024a). No natural brood stock has been available from Pond Spring since 2012 despite intensive sampling efforts from 2013 to 2015 (Kuhajda and Mitchell 2019). Benedict Spring (type locality) has dried completely at least 7 times since 2006; BTMs were rescued in each of these years and held in captivity until water levels were stable and then returned (Kuhajda and Mitchell 2019, USFWS

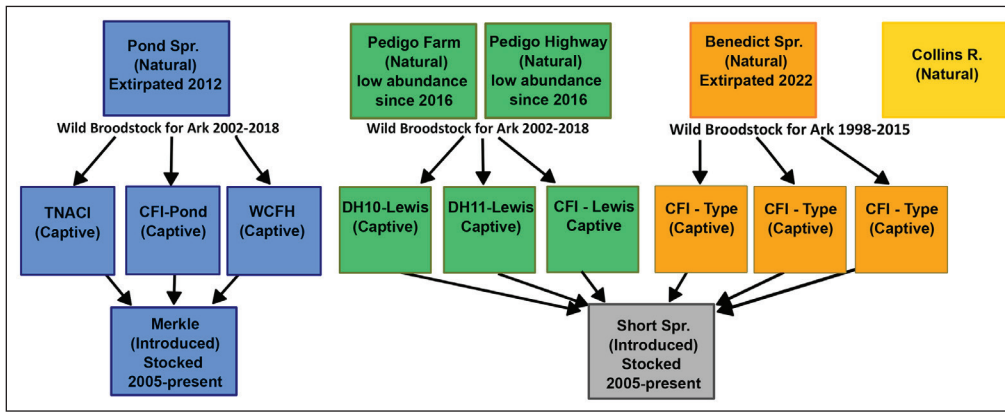

Figure 2. Flow chart illustrating the history of natural, captive, and introduced populations represented in this study.

2024b). In 2022, the spring completely dried, and Benedict Spring is currently not inhabited by BTMs (USFWS 2024b). Natural brood stock was still available from the Witty Creek MU after an extreme drought in 2016, but abundance was low (USFWS 2024c), and the last wild brood stock was collected in 2018 (P. Rakes, CFI, Knoxville, TN, pers. comm.).

An important goal for reintroduction programs is that captive-bred populations maintain the genetic diversity that was present within source populations while also maintaining adaptive differences between historically isolated populations. Captive populations often experience a reduction in genetic diversity due to both founder effects and small population sizes, reducing the species' long-term adaptive potential (Fraser 2008, George et al. 2009). Differences in selective pressures between captive and wild environments can drive adaptation to captivity, which may reduce fitness when hatchery fish are stocked in the wild (Frankham 2008). Genetic monitoring is a valuable tool in managing the genetic health of captive breeding programs, and is part of the propagation plan for the BTM (USFWS 2024b). Information on the genetic composition of captive populations can be used to design breeding strategies that minimize inbreeding and genetic drift, as well as to identify genes underlying phenotypic adaptations to hatchery conditions.

The objective of this project was to characterize genetic variation within and between source populations and captive ark populations of BTM at TNACI, CFI, DHFH, and WCFH to inform continued reintroduction efforts. Specifically, we sought to estimate loss of genetic diversity and changes in allelic composition in captive BTM populations that may have resulted from long-term isolation and limited population sizes. We generated genome-wide SNP genotypes for 4 native populations and 10 captive-bred populations. In addition, we used both mitochondrial sequence data and SNP genotypes to characterize genetic variation in the newly discovered BTM population in the upper Collins River. Results from this study will be used to inform the continued maintenance of captive populations of BTM that will serve as a source for future reintroductions. This project will also serve as a pilot study for the use of SNP-based protocols for long-term genetic monitoring of natural, captive, and reintroduced BTM populations.

Methods

Tissue samples

We sampled a total of 15 BTM populations for this study, including 4 natural populations, 1 introduced population, and 10 captive populations (Table 1). With a sterilized blade, we removed fin clips (2–3 mm^2), to be used for DNA extractions, from the caudal fin of each fish. We released fish back into the population following a brief recovery period. Tissue samples from a recently discovered natural population in the upper Collins River were collected by collaborators at the USFWS in 2023. We obtained DNA samples from natural populations at Pedigo Farm, Pedigo Highway, Pond Spring, and an introduced population at Short Spring (Benedict/Type and Pedigo Farm brood stock) from stored DNA samples used in a previous genetic survey of BTM (Hurt et al. 2017). Unfortunately, DNA was not available

for Benedict Spring at the time of this study. Tissues from natural populations were obtained during routine monitoring by collaborators at TNACI in Chattanooga, TN, except for fin clips from Pond Spring fish, which were collected by USFWS in 2011 prior to the extirpation of this population (Kuhajda and Mitchell 2019, Kuhajda et al. 2014, Zuber and Mattingly 2012). We acquired tissue samples from captive populations from 4 BTM populations at DHFH (DH3-Type, DH4-Type, DH10-Lewis, and DH11-Lewis), 1 population at TNACI, 1 population at WCFH, and 4 populations at CFI (CFI-Type, CFI-Lewis, CFI-Pond, and CFI-Merkle). We stored tissues in RNAlater Stabilization Solution (Sigma Aldrich, St. Louis, MO) or in 95% ethanol at -20 °C prior to extraction of DNA. We isolated DNA from fin clips with the Omega Bio-tek E.Z.N.A Blood and Tissue kit (Norcross, GA) using the manufacturer's protocol, except that the final elution was in water.

Mitochondrial sequencing and analyses

We amplified a ~700-bp region of the mitochondrial control region from 8 sampled individuals from the upper Collins River. We used PCR amplification primers and protocols following Hurt et al. (2017). We sequenced PCR products in both directions using Sanger sequencing on an ABI 3730XL automated sequencer (MCLAB, South San Francisco,CA). We imported and visualized sequence chromatograms using SEQUENCHER version 5.2 (Gene Codes Corp., Ann Arbor, MI). We exported consensus sequences to the program Bioedit (Hall 1999) and added them to an existing sequence alignment that included all unique BTM D-loop haplotypes identified by Hurt et al. (2017) and outgroup sequences obtained from GenBank.

Table 1. Sample sizes and standard measures of genetic diversity for 15 Barrens Topminnow populations included in GBS library construction. Population type (i.e. natural, introduced, or captive) and source populations are indicated for introduced and captive populations. *n* indicates the number of individuals included in the final SNP genotype file after filtering for low coverage individuals. Genetic diversity summary statistics include observed heterozygosity (H_o), expected heterozygosity (H_e), and the proportion of polymorphic SNPs (*P*).

Population	*n*	Pop. type	Source	H_o	H_e	*P*
Collins R.	16	Natural	-	0.041	0.052	0.778
Pedigo F.	14	Natural	-	0.033	0.053	0.607
Pedigo H.	16	Natural	-	0.033	0.050	0.647
Pond Spr.	20	Natural	-	0.050	0.055	0.600
Short Spr.	16	Introduced	Benedict and Lewis	0.038	0.057	0.636
CFI-Type	16	Captive	Benedict Spring	0.040	0.032	0.539
CFI-Lewis	19	Captive	Lewis	0.042	0.043	0.684
CFI–Pond	24	Captive	Pond Spring	0.047	0.042	0.549
CFI-Merkle	24	Captive	Merkle Spring	0.045	0.039	0.537
DH10–Lewis	10	Captive	Lewis	0.041	0.044	0.537
DH11-Lewis	10	Captive	Lewis	0.047	0.039	0.553
DH3-Type	6	Captive	Benedict Spring	0.032	0.026	0.537
DH4-Type	10	Captive	Benedict Spring	0.036	0.034	0.539
TNACI-Pond	20	Captive	Pond Spring	0.046	0.046	0.613
WCFH-Pond	12	Captive	Pond Spring	0.039	0.065	0.500

GBS Library preparation and sequencing

A total of 233 individuals were included for GBS sequencing. We quantified genomic DNA using the Quant-it Picogreen dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA) and standardized all samples to 8.5 ng/ul. Extracted DNA was digested with the restriction enzyme ApeKI. We ligated adaptors containing PCR binding sites and individual barcodes onto digested DNA. We pooled barcoded DNA and amplified PCR using primers that bind to the ligated adaptors (see Elshire et al. 2011 for primer sequences). We cleaned the resulting PCR products first using the Qiagen PCR purification kit (Hilden, Germany) and then again using the AxyPrep Mag PCR Clean-up kit (Axygen, Big Flats, NY). We determined the distribution of the fragment size in the PCR product by agarose gel electrophoresis. We sequenced barcoded libraries using the Illumina NextSeq 1000 (San Diego, CA) with P2 100-bp single-end read chemistry.

SNP discovery and filtering

We used 'process-radtags' in the Stacks program v. 2.68 to filter and demultiplex raw reads based on barcoded sequences (Catchen et al. 2011). We used the 7-step de novo clustering pipeline ipyrad v. 3.5 (Eaton 2014) to generate and filter SNP datasets used in downstream analyses. Quality filtering of raw sequence reads converted bases with Phred scores <33 to Ns; we removed reads with more than 5 Ns. We clustered reads using a sequence similarity threshold of 90% both within and between sampled individuals, with a minimum read depth of 6. We excluded individuals with fewer than 500,000 reads from downstream analyses. We removed loci with observed heterozygosity (H_o) greater than 0.5 to filter out possible paralogs. We then filtered the final SNP dataset to remove loci deviating from Hardy–Weinberg equilibrium ($P < 0.05$), loci genotyped in less than 60% of individuals, and SNPs with a minor allele frequency less than 0.01. Only 1 SNP per tag was retained per locus.

SNP summary statistics

We estimated standard measures of genetic diversity including H_o and expected heterozygosity (H_e) values and the proportion of polymorphic loci (P), as well as pairwise F_{ST} and Nei's standard pairwise genetic distances (Weir and Cockerham 1984, Nei 1987), using the R package 'Hierfstat' v. 0.5-11 (Goudet et al. 2015). We used nonparametric bootstrapping (100 reps) on pairwise F_{ST} estimates to generate 95% confidence intervals using the 'boot.ppfst' function in 'Hierfstat'. We used the R package 'poppr' v. 2.9.6 to construct a neighbor-joining tree based on Nei's DA pairwise genetic distances (Kamvar et al. 2014). We assessed support at nodes by nonparametric bootstrap (1000 replicates) using the 'aboot' function in 'poppr'.

Population assignment analyses

We estimated the optimal number of genetic clusters (K) based on genomic SNPs using 2 different methods including a Bayesian-based assignment test and multivariate analysis. We performed Bayesian assignment tests using the program STRUCTURE 2.3.4 (Pritchard et al. 2003). Values of K (number of genetic populations) varied from 1 to 12 populations with 10 replicate runs per

value of K; we performed MCMC simulations for a burn-in of 50,000 iterations and retained an additional 1 x 10^6 iterations for the final analysis. We examined the optimal number of populations (K) using the Delta K method (Evanno et al. 2005). Results were summarized using the software package 'CLUMPAK' (Kopelman et al. 2015).

We performed discriminant analysis of principal component (DAPC) using the 'adegenet' package in R (v. 2.1.10; Jombart and Ahmed 2011, Jombart and Collins 2015). We first used the 'find.clusters' function to identify the optimal value of K based on a Bayesian information criterion (BIC) process and then used the optimal K to perform a DAPC analysis to describe the relationship between the genetic clusters. We assessed individual membership probabilities using the function 'compoplot'.

Results

Mitochondrial sequence analysis

Sanger sequencing of the mitochondrial D-loop resulted in a ~695-bp sequence alignment. All 8 sequenced individuals from the upper Collins River population shared a single mitochondrial haplotype that contained 1 fixed substitution unique to this population (Fig. 3).

SNP bioinformatics and summary statistics

The average number of retained sequence reads per individual generated from sequencing of GBS libraries was 3,810,070. A total of 233 individuals were retained

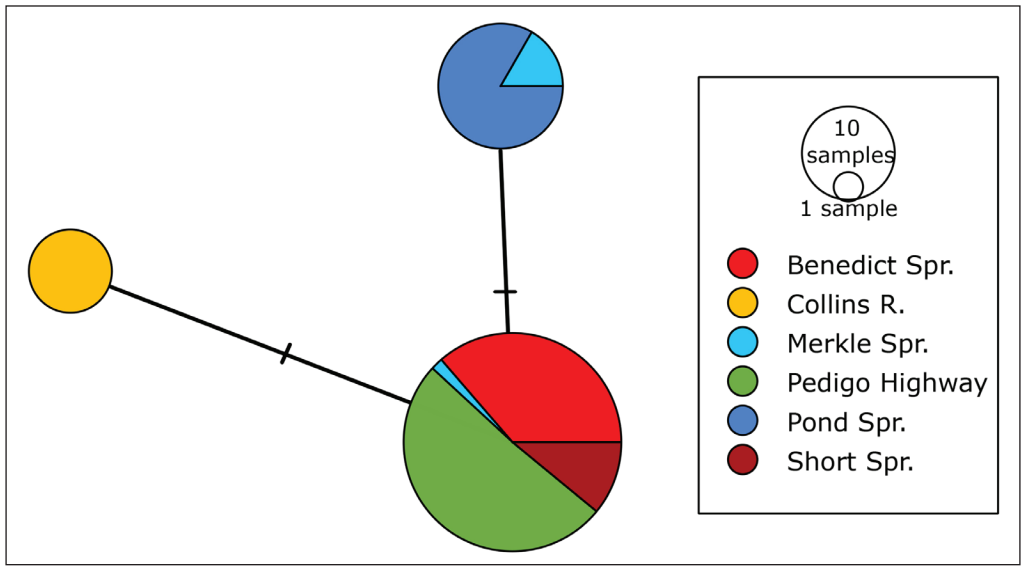

Figure 3. Minimum spanning haplotype network based on 695 bp mitochondrial D-loop sequences from 6 populations of Barrens Topminnow. Each circle represents a unique haplotype, with its size proportional to frequency. Lines connecting haplotypes indicate mutational steps. Colors correspond to different populations as indicated in the legend.

after removal of low-coverage individuals (defined as <500 total reads). A total of 1,010,622 loci were recovered from the de novo assembly in ipyrad. Additional filtering for low-coverage SNPs, Hardy–Weinberg equilibrium, and 1 SNP per tag retained 1888 SNPs for the final dataset. Estimates of expected heterozygosity were similar across the 4 sampled natural populations, averaging 0.053, and the proportion of polymorphic loci was 66%. Estimates of within-population genetic diversity were higher for natural populations than for captive populations. Average expected heterozygosity across the 10 captive populations was 0.042 and the proportion of polymorphic loci was 57% (Table 1).

Pairwise genetic-distance estimates indicated that the upper Collins River population was genetically distinct from all other natural populations of BTM. The average pairwise F_{ST} value for comparisons between upper Collins River and the other 3 natural populations (Pedigo Farm, Pedigo Highway, and Pond Spring) was 0.294 (Fig. 4). This value is similar to pairwise F_{ST} values for comparisons between Pond Spring and other natural populations, which averaged 0.297. Genetic distances between natural Pond Spring and captive Pond Spring populations varied from 0.076 (CFI-Pond) to 0.128 (TNACI), averaging 0.104. Genetic distances between natural Pedigo F./Pedigo H. and captive Lewis populations averaged 0.075 (DH11-Lewis) and 0.099 (CFI-Lewis). Benedict Spring individuals (type locality) were not included in the analysis. Pairwise F_{ST} estimates for comparisons between captive Benedict Spring populations were very low: F_{ST} averaged only 0.013

The neighbor-joining tree recovered 3 monophyletic clades corresponding to the 3 genetic lineages maintained in the captive-breeding program (Fig. 5). All captive

Sites		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Collins R.	–														
2	Pedigo F.	**0.266**	–													
3	Pedigo H.	**0.318**	**0.100**	–												
4	Pond Spr.	**0.300**	**0.275**	**0.317**	–											
5	Short Spr.	**0.267**	**0.100**	**0.170**	**0.276**	–										
6	CFI-Type	**0.362**	**0.300**	**0.343**	**0.396**	**0.127**	–									
7	CFI-Lewis	**0.289**	**0.076**	**0.122**	**0.334**	**0.135**	**0.335**	–								
8	CFI–Pond	**0.292**	**0.263**	**0.307**	**0.076**	**0.238**	**0.330**	**0.283**	–							
9	CFI-Merkle	**0.308**	**0.274**	**0.322**	**0.087**	**0.258**	**0.392**	**0.307**	**0.047**	–						
10	DH10-Lewis	**0.264**	**0.085**	**0.132**	**0.290**	*0.020*	**0.105**	**0.106**	**0.249**	**0.297**	–					
11	DH11-Lewis	**0.280**	**0.063**	**0.087**	**0.333**	**0.111**	**0.348**	*-0.006*	**0.301**	**0.331**	**0.100**	–				
12	DH3-Type	**0.322**	**0.234**	**0.297**	**0.350**	**0.069**	*0.002*	**0.319**	**0.309**	**0.372**	**0.082**	**0.341**	–			
13	DH4-Type	**0.315**	**0.216**	**0.276**	**0.345**	**0.071**	*0.010*	**0.272**	**0.295**	**0.358**	**0.045**	**0.277**	*0.000*	–		
14	TNACI-Pond	**0.272**	**0.215**	**0.269**	**0.128**	**0.144**	**0.224**	**0.263**	**0.072**	**0.108**	**0.159**	**0.258**	**0.178**	**0.169**	–	
15	WCFH-Pond	**0.265**	**0.171**	**0.249**	**0.108**	**0.112**	**0.248**	**0.243**	**0.076**	**0.126**	**0.144**	**0.233**	**0.177**	**0.195**	**0.070**	–

≥ 0.35
0.20–0.35
0.10–0.20
≤ 0.05

Figure 4. Pairwise F_{ST} estimates based on 1888 SNPs from 15 sampled populations of Barrens Topminnow. Statistical significance was determined using 95% confidence intervals. Pairwise comparisons significantly different from zero are shown in bold font and nonsignificant comparisons are shown in italics.

populations clustered with their initial source populations except for DH10-Lewis and DH11-Lewis. These 2 populations were recovered as outgroups to the Benedict Spring clade, but were initially founded by Lewis Creek broodstock. Short Spring (a mixture of Benedict and Lewis populations) was basal to the Lewis Creek clade. Lastly, the upper Collins River population was genetically distinct from all other sampled populations.

Population assignments

Results from Bayesian assignment tests indicated that K = 5 populations provided the best explanation for genetic variation at SNP genotypes based on both the Delta K and Ln(X|K) methods (Fig. 6). Based on examination of barplots, we assigned individuals to 4 major genetic clusters, and additional values of *K* appeared as admixture across multiple populations. Genetic cluster 1 was exclusively represented by all individuals sampled from the natural population in the upper Collins River. Genetic cluster 2 represented fish from the Lewis Creek lineage, which included populations from natural sites at Pedigo Farm, Pedigo Highway, and captive populations at DHFH (some DH10-Lewis individuals and all DH11-Lewis) and CFI-Lewis. The introduced population at Short Spring demonstrated admixture between genetic clusters 2 and 4. Genetic cluster 3 included samples from the natural population at Pond Spring and captive populations at TNACI, WCFH, CFI-Pond, and CFI-Merkle. Captive populations at TNACI and WCFH demonstrated some admixture with cluster 4. Captive populations from Dale Hollow (DH3-Type, some DH10-Lewis, and DH4-Type) and CFI-Type were assigned

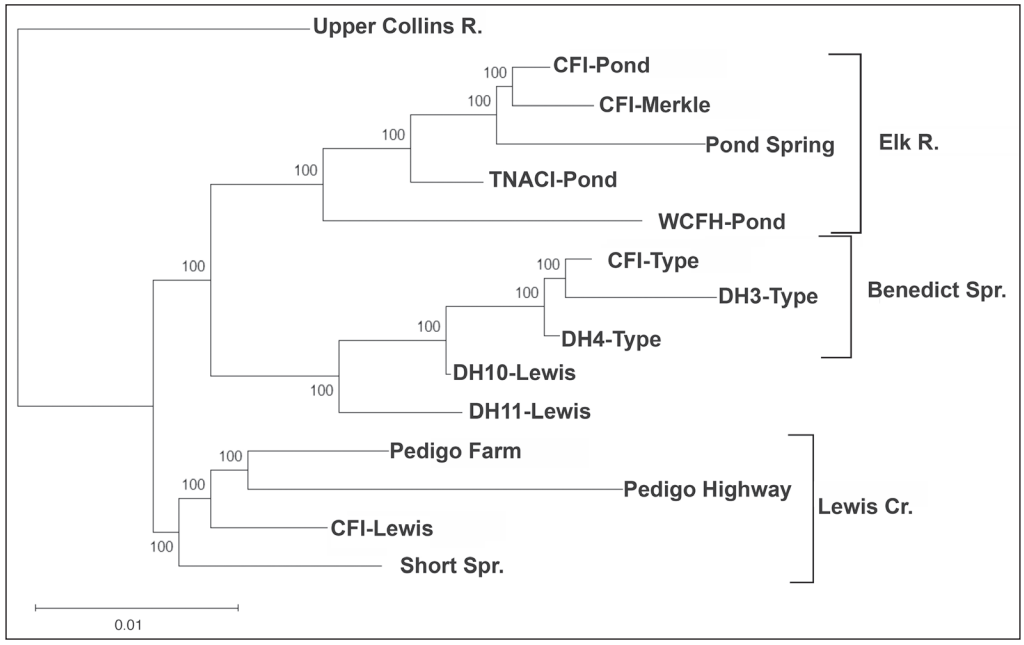

Figure 5. Neighbor-joining tree based on Nei's standard genetic distance estimates from 1888 SNP loci for 15 sampled populations of Barrens Topminnow. Nodal support is based on 1000 bootstrap replicates.

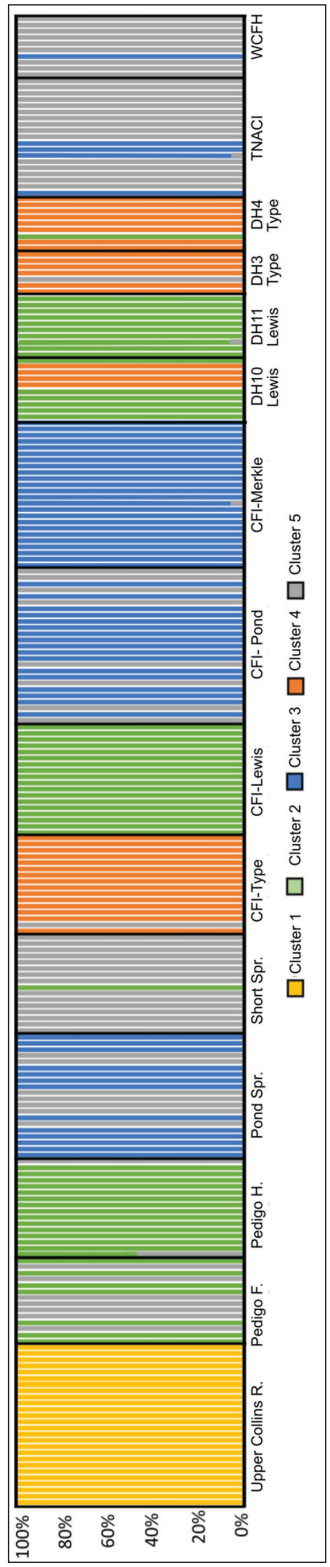

Figure 6. Results of Bayesian assignment test based on 1888 SNPs generated from GBS sequencing for K= 5 populations. Barplots indicate individual assignment probabilities for samples from 15 sampled populations of Barrens Topminnow.

to cluster 4. Admixture of cluster 4 also appeared in captive populations at TNACI, WCFH, and CFI-Pond. Genetic cluster 5 appeared as a low-level admixture across multiple populations.

Results from the multivariate analysis (DAPC) of SNP genotypes were generally consistent with results from Bayesian assignment tests. BIC values indicated K = 5 as the optimal number of genetic clusters (Fig. 7). Membership probability plots from DAPC analysis have been included (see Supplemental Figure 1 in Supplemental File 1, available online at https://www.eaglehill.us/SENAonline/suppl-files/s24-sp13-S2934b-Hurt-s1, and for BioOne subscribers, at https://www.doi.org/10.1656/S2934b.s1) to inform individual assignments to genetic clusters; these plots show the probability of each individual's assignment to clusters based on the retained discriminant functions. The plots demonstrated sharing of genetic clusters (clusters 3 and 4) across multiple populations, which was not observed in Bayesian assignment results. DAPC cluster 1 was exclusively represented by individuals sampled from the upper Collins River. Cluster 2 appeared in subsets of individuals from natural populations at Pedigo Farm and Pedigo Highway as well as captive populations from the Lewis Creek lineage (CFI-Lewis, DH10-Lewis, and DH11-Lewis). Cluster 3 was assigned to individuals from Pond Spring and captive populations at CFI-Pond, CFI-Merkle, TNACI, and WCFH. Cluster 4 was represented by captive populations (Benedict/Type-lineage) at CFI and DHFH (DH3-Type, DH4-Type, and DH10-Lewis). Lastly, cluster 5 was assigned to a subset of individuals from natural populations at Pedigo Farm, Pedigo Highway, and Pond Spring and from captive populations at DHFH (DH4-Type, DH10-Lewis, DH11-Lewis) and CFI-Pond.

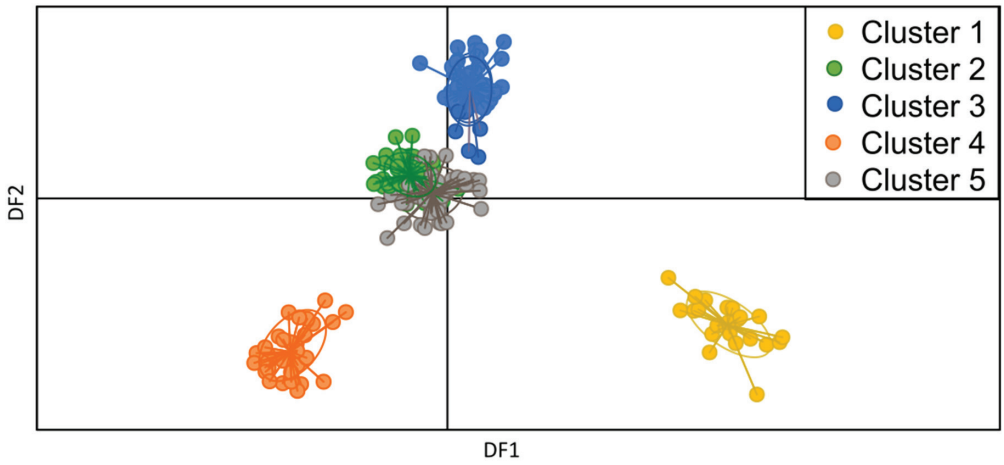

Figure 7. Discriminant analysis of principal components (DAPC) scatter plot based on 1888 SNPs from 15 populations of Barrens Topminnow. The optimal number of clusters (K = 5) was determined using Bayesian information criterion (BIC). Axes represent the first 2 discriminant functions, which maximize separation among clusters.

Discussion and Conclusions

This study investigated the genetic consequences of management strategies for BTM that included long-term captive breeding and reintroductions. Analysis of 1888 genomic SNPS from natural BTM populations at Pedigo Farm, Pedigo Highway, and Pond Spring are consistent with previous studies based on microsatellite loci that suggested the presence of multiple, evolutionarily distinct lineages (Hurt et al. 2017). Results from our study also provide valuable information on the genetic composition of the newly discovered BTM population in the upper Collins River. Captive populations of BTM were assigned to the same genetic cluster as their corresponding source population; however, there was evidence of a loss of genetic variation and drift that is often seen in captive-bred populations. Based on these findings, we provide specific management suggestions that may improve the long-term adaptive potential of this species.

Population structure

Results from the analysis of population structure based on SNP markers is consistent with patterns observed from an earlier survey using microsatellite markers (Hurt et al. 2017). Assignment results from both sets of markers identified natural populations at Pond Spring (cluster 3; Figs. 6, 7) and Pedigo Farm/Pedigo Highway (Lewis populations, cluster 2) as distinct genetic lineages with high probability. The previous microsatellite study identified Benedict Spring (type locality) as a third genetic lineage. Samples from Benedict Spring were not available for the present study. However, Benedict Spring is represented here by captive populations at DHFH (DH3-Type and DH4-Type) and at CFI (CFI-Type); these populations were initiated from Benedict Spring brood stock and were recovered as separate genetic entities in both Bayesian assignment tests and DAPC (cluster 4). Assignment results indicated that there may have been mixing of stocks among captive BTM populations at DHFH, where both Benedict Spring and Lewis populations are held. One individual from DH4-Type was assigned to cluster 2 (Lewis), and 4 individuals from DH10-Lewis were assigned to cluster 4 (Benedict Spring). Finally, the introduced population at Short Spring showed evidence of population admixture between cluster 2 and cluster 4, consistent with earlier microsatellite results (Hurt et al. 2017). Stocking records indicate that Short Spring was founded with a mixture of Benedict Spring and Pedigo Farm brood stock. Short Spring individuals have retained genetic variation from both of these source populations based on results from our Bayesian assignment tests (Fig. 6).

Results from both mitochondrial sequences and SNP genotypes indicated that the newly discovered BTM population in the upper Collins River is a valuable genetic resource for the conservation of this species. Although the Collins River mitochondrial haplotype only possessed a single base pair difference from other BTM haplotypes, this difference is significant as BTM is known to have very little genetic variation in the mitochondrial genome. A previous survey of mtDNA sequence diversity across 8 BTM populations (60 individuals) identified 1 indel and 1 polymorphic site for a total of 3 haplotypes (Hurt et al. 2017). Genetic

differentiation of the upper Collins River population was more pronounced in results from our survey of genomic SNP genotypes. Both the Bayesian assignment tests and DAPC assigned upper Collins River individuals to their own unique genetic cluster (cluster 1). The average pairwise F_{ST} for the upper Collins River samples was similar to pairwise F_{ST} values for the Pond Spring population, which was assigned as a distinct ESU based on evidence from microsatellites (Hurt et al. 2017, 2019). Collectively, our results warrant the designation of the upper Collins River population as its own distinct ESU that should be managed separately from other natural populations of BTM.

Captive populations

The continued persistence of BTM relies heavily on the maintenance of captive populations that are used for stocking BTM into suitable habitats throughout their historical range (Kuhajda et al. 2014, USFWS 2024b). The breeding program for BTM was designed to keep captive populations from different natural sources separated. Assignment tests and DAPC results showed that captive populations mostly shared the same genetic cluster as their source population; however, a measurable loss of genetic variation was detected. Nevertheless, we did find evidence of mixed stocks at DHFS, where both Benedict Spring and Pedigo Farm brood stock have been maintained and were recovered in the same breeding population. All captive populations showed some evidence of genetic drift from their natural source population as revealed by pairwise F_{ST} estimates, which measures shifts in allele frequencies. Comparisons between the BTM population at TNACI and its source population, Pond Spring, had a pairwise F_{ST} of 0.128, which can be interpreted as moderate genetic differentiation. For reference, the guidelines often used to interpret F_{ST} is that values less than 0.05 indicate insignificant differentiation, 0.05–0.15 represents moderate differentiation, 0.15–0.25 suggests great genetic differentiation, and values >0.25 indicate very great genetic differentiation (Hartl and Clark 1997). Interestingly, populations founded with Benedict Spring stock showed little evidence of drift as indicated by low pairwise F_{ST} values (all <0.05). In the microsatellite survey, Benedict Spring had the least amount of genetic variation and was fixed for a single allele at 7 out of the 14 surveyed loci. It is likely that Benedict Spring is also fixed for many of the SNP loci surveyed here, limiting potential shifts in allele frequencies. This low amount of genetic variation is likely the result of the Benedict Spring completely drying at least 7 times since 2006, repeatedly creating a bottleneck as only a portion of the population was rescued in each of these years (Kuhajda and Mitchell 2019, USFWS 2024b).

Estimates of genetic variation indicated that captive BTM populations have a reduction of genetic variation when compared to their natural source populations. Expected heterozygosity from captive populations from Lewis Creek brood stock were on average 21% lower than estimates from Pedigo Creek and Pedigo Farm. For captive populations from Pond Spring brood stock at CFI and TNACI, the average reduction of expected heterozygosity is 23%. Interestingly, WCFH had very high estimates of H_e that greatly exceeded estimates of H_o. Several factors

can inflate estimated H_e relative to H_o including hidden population structure, nonrandom mating, and bioinformatic artifacts from SNP processing. Some loss of genetic variation in captivity is unavoidable, and currently there are no established guidelines for how much genetic variation should be maintained for ensuring species viability. A loss of 1% per generation is seen as acceptable for agricultural lines and in livestock breeding programs (Frankel and Soulé 1981, Naish and Haard 2008). In general, captive-breeding programs should aim to retain as much genetic variation as possible in order to ensure long-term fitness and adaptive potential (Fraser et al. 2008, George et al. 2009).

Management recommendations

Results from this study support the MU and ESU designations introduced by Hurt et al. (2017). In addition, we recommend that the upper Collins River population be managed as a third, distinct ESU. Establishment of a separate captive and/or introduced wild population by upper Collins River brood stock should be considered in order to protect this genetic resource. As outlined in the Barrens Topminnow Propagation Plan (USFWS 2024c) and in the Recovery Plan for Barrens Topminnow (USFWS 2024b), management of captive populations should establish hatchery protocols that minimize the degree of kin-mating while maintaining the genetic integrity of distinct evolutionary lineages. Exchange of individuals between captive and reintroduced populations established from brood stock belonging to the same ESU/MU could effectively reduce the amount of inbreeding and improve fitness.

Future directions

Captive-breeding programs are increasingly being used as a last resort to prevent extinction of at-risk species. However, reintroductions of captive-reared fish have had mixed success, which can often be attributed to adaptation to captive conditions and loss of overall genetic diversity. In captive-reared salmon, reproductive capacity was reduced by 40% when reintroduced into the wild (Araki et al. 2007). Surveys of BTM populations have also found that hatchery-raised BTM in stocked populations have lower survival and recruitment than BTM in natural populations (Ennen et al. 2021), but the presence of invasive mosquitofish in most stocked populations is a major factor for their lower fitness. The application of genome-wide SNP datasets and molecular tools such as transcriptome sequencing can improve the success of reintroduction programs by informing breeding programs about adaptive genetic diversity that directly impacts fitness. Identifying functional regions of the genome that correlate with fitness measures can help pinpoint adaptive alleles. Information about functionally important genetic variation can then be used to enhance breeding programs, improve reproductive fitness, and increase survival rates in the wild.

Acknowledgments

Thanks to Warren Stiles, Matt Hamilton, Pat Rakes, CFI, DHFH, TNACI, and WCFH for providing tissue samples. The illustration of the male BTM was provided by Joe Tomelleri.

The map of sampled BTM populations was provided by Kurt Snider. Funding for this project was provided by the US Fish and Wildlife Service.

Literature Cited

Araki, H., B. Cooper, and M.S. Blouin. 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318(5847):100–103.

Catchen, J.M., A. Amores, P. Hohenlohe, W. Cresko, and J.H. Postlethwait. 2011. Stacks: Building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics, 1(3):171–182.

Eaton, D.A. 2014. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30(13):1844–1849.

Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high-diversity species. PLoS ONE 6(5):e19379.

Ennen, J.R., B.R. Kuhajda, S. Fix, S.C. Sweat, B. Zuber, A.V. Watts, H.T. Mattingly, and K.K. Cecala. 2021. Assessing the success of conservation efforts for a North American topminnow at risk of extinction from spatially variable mosquitofish invasions. Freshwater Biology 66(3):458–467.

Frankel, O.H., and M.E. Soulé. 1981. Conservation and Evolution. Sinauer Associates, Sunderland, MA. 327 pp.

Frankham, R. 2008. Genetic adaptation to captivity in species conservation programs. Molecular Ecology 17(1):325–333.

Fraser, D.J. 2008. How well can captive-breeding programs conserve biodiversity? A review of salmonids. Evolutionary Applications 1(4):535–586.

George, A.L., B.R. Kuhajda, J.D. Williams, M. Cantrell, P.L. Rakes, and J.R. Shute. 2009. Guidelines for propagation and translocation for freshwater fish conservation. Fisheries 34:529–545.

Goudet, J., T. Jombart, and M. J. Goudet. 2015. Hierfstat: Estimation and tests of hierarchical F-statistics. Version 0.5-11. R package. Available online at https://cran.r-project.org/web/packages/hierfstat/index.html.

Hall, T.A. 1999. BioEdit: A user–friendly biological sequence-alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41(41):95–98.

Hartl, D.L., and A.G. Clark. 1997. Principles of Population Genetics. 3rd Edition. Sinauer Associates, Sunderland, MA. 481 pp.

Hurt, C., B. Kuhajda, A. Harman, N. Ellis, and M. Nalan, 2017. Genetic diversity and population structure in the Barrens Topminnow (*Fundulus julisia)*: Implications for conservation and management of a critically endangered species. Conservation Genetics 18:1347–1358.

Hurt, C., N. Ellis, A. Harman, and C. Savage. 2019. Genetic structure at the major histocompatibility complex in the endangered Barrens Topminnow (*Fundulus julisia*). Southeastern Naturalist 18(1):19–36.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3–1: New tools for the analysis of genome-wide SNP data. Bioinformatics 27(21):3070–3071.

Jombart, T., and C. Collins, 2015. Analysing genome-wide SNP data using adegenet 2.0. 0. Imperial Colleege London–MRC Centre for Outbreak Analysis and Modeling, London, UK. 48 pp.

Kamvar, Z. N., J. F. Tabima, and N. J. Gronwald. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. Molecular Ecology Resources 14(3):413–420.

Kopelman, N.M., J. Mayzel, M. Jakobsson, N.A. Rosenberg, and I. Mayrose. 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15(5):1179–1191.

Kuhajda, B., and S. Mitchell. 2019. Population status and monitoring of the imperiled Barrens Topminnow, *Fundulus julisia*; 2014–2015 monitoring report. US Fish and Wildlife Service, Cookeville TN. 139 pp.

Kuhajda, B., D. Neely, and K. Alford. 2014. Population status and monitoring of the imperiled Barrens Topminnow, *Fundulus julisia*; 2013 monitoring report. Report to US Fish and Wildlife Service, Cookeville, TN. 122 pp.

Laha, M., and H.T. Mattingly. 2007. Ex situ evaluation of impacts of invasive mosquitofish on the imperiled Barrens Topminnow. Environmental Biology of Fishes, 78(1):1–11.

Moritz, C., 1994. Defining "evolutionarily significant units" for conservation. Trends in Ecology and Evolution 9(10):373–375.

Naish, K.A., and J.J. Hard. 2008. Bridging the gap between the genotype and the phenotype: Linking genetic variation, selection, and adaptation in fishes. Fish and Fisheries 9(4):396–422.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press. 512 pages.

US Fish and Wildlife Service (USFWS). 2024a. Barrens Topminnow (*Fundulus julisia*) Species Status Assessment, Version 1.2. Washington, DC. 46 pp. Available online at https://iris.fws.gov/APPS/ServCat/DownloadFile/266648.

USFWS. 2024b. Recovery plan for Barrens Topminnow (*Fundulus julisia*). Southeast Region, Atlanta, GA. 16 pp.

USFWS. 2024c. Barrens Topminnow (*Fundulus julisia*) propagation plan. Southeast Region, Atlanta, GA. pp. 13 pp.

Weir, B.S., and C.C. Cockerham, 1984. Estimating *F*-Statistics for the analysis of population structure. Evolution 38:1358–1370.

Zuber, B., and H.T. Mattingly. 2012. Population status and monitoring of the imperiled Barrens Topminnow, *Fundulus julisia*: Annual topminnow monitoring report. Report to US Fish and Wildlife Service, Cookeville, TN. 36 pp.