Zoological Journal of the Linnean Society, 2017, **180**, 461–474. With 4 figures.

Species delimitation and phylogeography of the studfish *Fundulus catenatus* species group (Ovalentaria: Cyprinodontiformes)

PETER J. HUNDT[1,2,3]*, PETER B. BERENDZEN[4] and ANDREW M. SIMONS[2,3]

[1]*Conservation Biology Graduate Program, University of Minnesota, 1987 Upper Buford Circle, St. Paul MN, 55108, USA*
[2]*Bell Museum of Natural History, University of Minnesota, 100 Ecology, 1987 Upper Buford Circle, St. Paul MN, 55108, USA*
[3]*Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, 1980 Folwell Avenue, St. Paul MN, 55108, USA*
[4]*Department of Biology, University of Northern Iowa, McCollum Science Hall, Cedar Falls IA, 50614-0421, USA*

Received 27 October 2015; revised 8 June 2016; accepted for publication 15 July 2016

The Central Highlands of North America have an old and complex geological and biological history, with extensive unexplored cryptic diversity. We examined the species limits of the *Fundulus catenatus* species group (*F. catenatus*, *Fundulus bifax*, and *Fundulus stellifer*) using two phylogenetic approaches. *Fundulus catenatus* was densely sampled from five geographically disjunct regions across the eastern USA. We sequenced four nuclear introns and used species tree analyses and two species delimitation methods to determine historical relationships and investigate species diversity within the group. Both methods strongly supported the presence of two reciprocally monophyletic species within *F. catenatus*. Species tree analysis of four nuclear introns (*stx5a, ncl1, rpsa, and rps3*) from the *F. catenatus* species group (including a tip for *F.* sp. cf. *catenatus*) recovered the two Mobile Basin species as sister to a strongly supported clade of *F. catenatus* and *F.* sp. cf. *catenatus*. Cytochrome *b* sequences were used in phylogeographical analyses of the two putative species. In one species we identified signatures of population expansion whereas the other exhibited genetic structure consistent with isolation of populations.

doi: 10.1111/zoj.12485

ADDITIONAL KEYWORDS: Central Highlands – cryptic species – Fundulidae – North American ichthyofauna – taxonomy.

INTRODUCTION

The identification and delimitation of species are critical components of systematic and evolutionary biology. The evolutionary species, a lineage of populations distinct from other lineages and with a unique evolutionary trajectory and fate (Wiley, 1981), is a fundamental unit of biodiversity. Species have traditionally been identified using morphological criteria, a reflection of earlier typological definitions as well as historical technologies that governed the types of data available. However, there is no reason to expect that species must be morphologically distinct, particularly under contemporary, evolutionary species concepts (Wiley, 1981). The advent of novel technologies, such as DNA sequencing, and the development of new analytical tools (reviewed in Carstens *et al.*, 2013), facilitate tests of species boundaries and ultimately aid species delimitation. It is expected that these will lead directly to an increased assessment of diversity as new morphologically cryptic species are discovered.

The North American ichthyofauna, one of the best-known ichthyofaunas of any continent, has been

*Corresponding author. E-mail: hundt002@umn.edu

under intense study for almost 200 years (Myer, 1964) but recent work has demonstrated the existence of morphologically cryptic species (Egge & Simons, 2006; Niemiller, Near & Fitzpatrick, 2012). Herein we investigate the existence of an undescribed species of studfish in the *Fundulus catenatus* species group (Ghedotti & Davis, 2013) and employ phylogeographical techniques to explore geographical boundaries and genetic signatures of expansion. The *F. catenatus* species group contains three described species: northern studfish, *F. catenatus* Storer, stippled studfish, *Fundulus bifax* Cashner & Rogers, and southern studfish, *Fundulus stellifer* Jordan (Figs 1, 2). *Fundulus catenatus* can be found in five disjunct regions in the eastern USA, separated by unsuitable lowland habitats: Ozark, Ouachita, and Eastern Highlands as well as small, isolated regions in eastern Indiana and south-western Mississippi (Fig. 2). In the Ozark highlands, northern studfish are found in tributaries to the Missouri River, direct tributaries to the Mississippi River, the upper Arkansas River, and the Black and White river drainages; in the Ouachita highlands, they occur in the Red, Ouachita, and Saline drainages; in the Eastern Highlands, they are widespread in the Tennessee, Green, Cumberland, and Kentucky drainages; in Indiana, they are restricted to small tributaries to the East Fork White River; and in south-west Mississippi they are restricted to tributaries of the Homochitto and Amite rivers (Fig. 2). The northern studfish is closely related to two taxa restricted to the Mobile Basin, the southern studfish, *F. stellifer*, and the stippled studfish *F. bifax* (Fig. 2). The systematics and morphology of the studfish group were reviewed by Ghedotti, Simons & Davis (2004) and revised by Ghedotti & Davis (2013).

The northern studfish exhibits minor morphological variation across its range and eight 'populations' were suggested based on a suite of meristic and morphological data (Thomerson, 1969). Thomerson (1969) speculated that this species initially dispersed west across the Mississippi River from the Eastern Highlands, suggesting that Ozark Highland *F. catenatus* are most closely related to Cumberland River *F. catenatus*, and Ouachita Highland *F. catenatus* are most closely related to Tennessee River *F. catenatus*. The Ozark Highland *F. catenatus* is morphologically homogeneous and occupies a large area, which Thomerson interpreted as indicative of recent dispersal (Fig. 2). Thomerson suggested that the disjunct southern Mississippi *F. catenatus* were the product of dispersal, specifically by tributary hopping down the Mississippi River.

The distribution of the *F. catenatus* species group was considered informative and consistent with the Central Highlands vicariance hypothesis (Mayden, 1988), a prediction that was tested by Grady, Cashner & Rogers (1990) and Strange & Burr (1997). Grady *et al.* (1990) used allozyme data to investigate relationships, and argued that these pointed to a different historical interpretation than the morphology. They observed an east–west split within *F. catenatus*, with Eastern Highland *F. catenatus* sister to Indiana, Ozark, Ouachita, and lower Mississippi *F. catenatus* (Fig. 2). Grady *et al.* (1990) interpreted their results as consistent with a widespread ancestral distribution subdivided by a series of vicariance

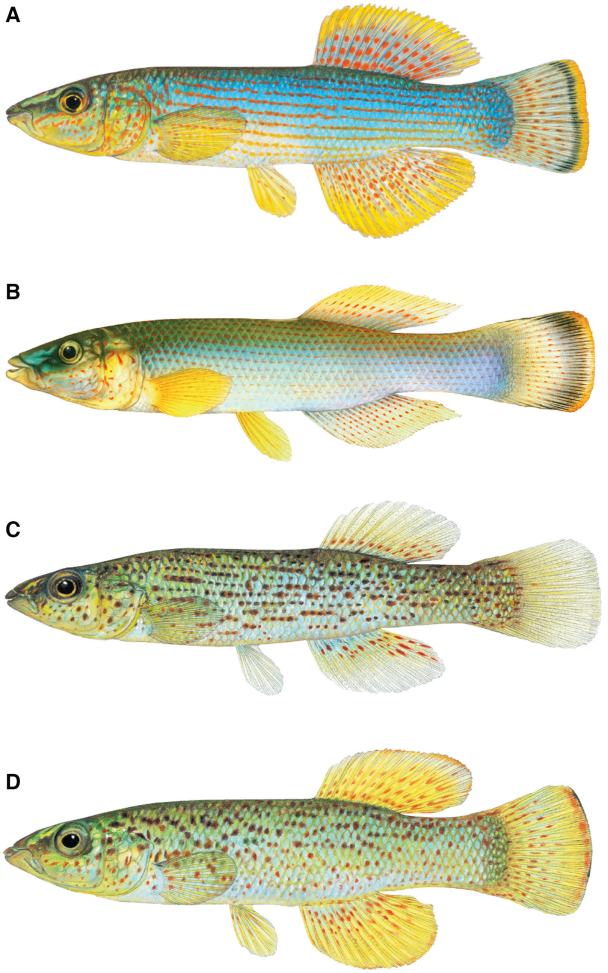

Figure 1. Illustrations of male specimens from the *Fundulus catenatus* species group: A, northern studfish, *Fundulus catenatus* (Paint Rock River, Tennessee River drainage, Jackson County, AL, UAIC 11147.05); B, *Fundulus* sp. cf. *catenatus* (Jack's Fork near Alley Springs, Shannon Co., MO, MHP 15157); C, stippled studfish, *Fundulus bifax* Cashner & Rogers (Gold Branch, Tallapoosa River drainage, Elmore County, AL, UAIC 10861.01); D, southern studfish, *Fundulus stellifer* Jordan (White Oak Creek, Alabama River drainage, Dallas County, AL, UAIC 10848.05). Illustrations by Joseph R. Tomelleri.

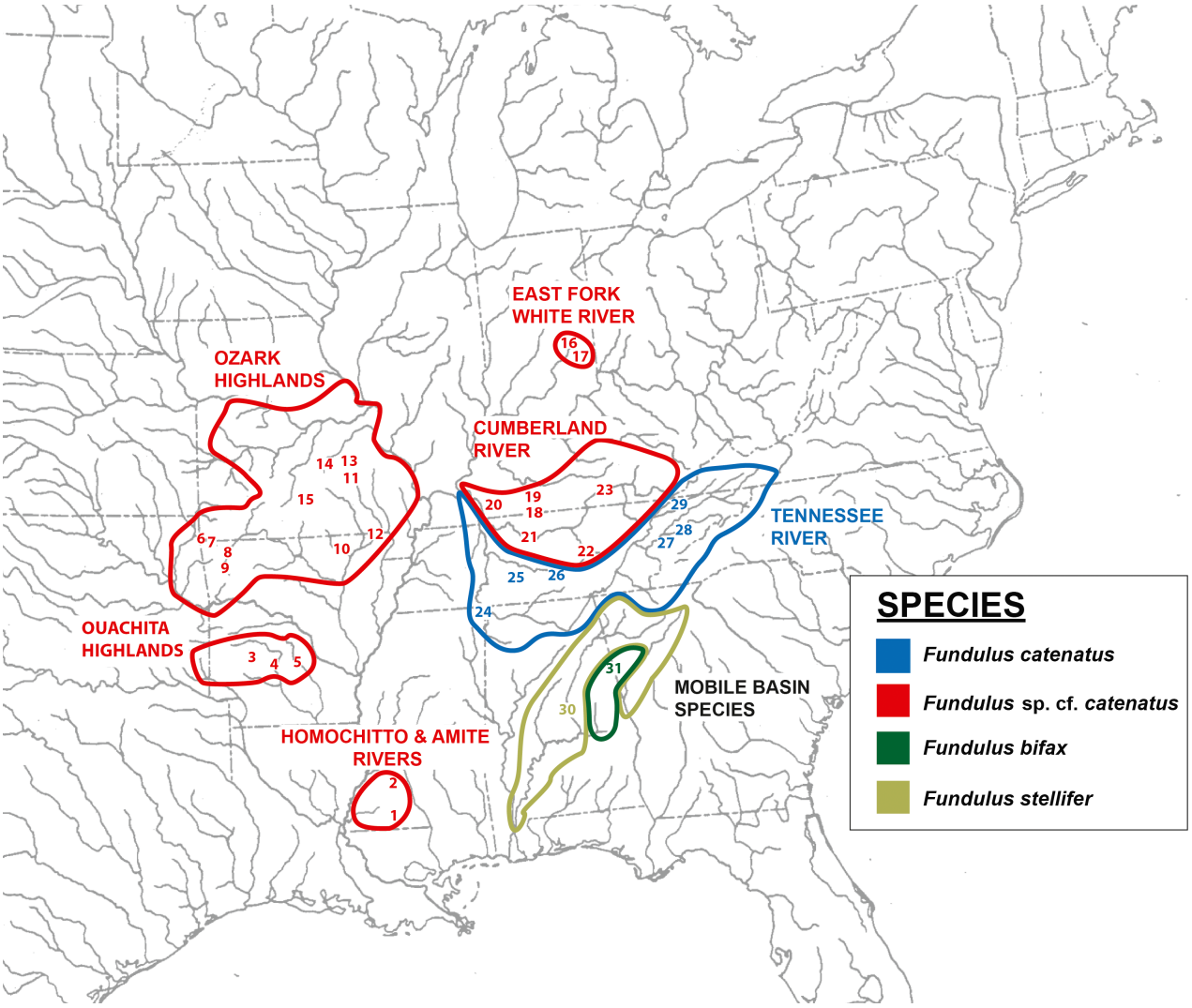

Figure 2. Map of south-eastern USA depicting the range and sampling locations of the *Fundulus catenatus* species group (see Table 2).

events. The southern Mississippi *F. catenatus* were thought to be isolated by vicariance or possibly established by dispersal via larval drift (Rogers & Cashner, 1987) or tributary hopping (Thomerson, 1969). Strange & Burr (1997) used mitochondrial restriction fragment length polymorphisms analysed with Dollo parsimony and examined the phylogeography of *F. catenatus* as part of a larger comparative test of the Central Highlands vicariance hypothesis. They reported a close relationship between the Indiana and Ozark *F. catenatus*, which were sister to most Eastern Highland *F. catenatus* (Tennessee and Cumberland rivers). They also identified paraphyly of the Mobile Basin species, with *F. stellifer* sister to *F. bifax* plus *F. catenatus*.

Phylogeographic analysis of *F. catenatus* using only mitochondrial *cytochrome b* (*cytb*) sequences (A. M. Simons, unpubl. data) indicated a deep division between Tennessee River *F. catenatus* and all remaining samples, including all other Eastern Highland *F. catenatus*. These data also indicated that *F. catenatus* was paraphyletic with respect to the Mobile Basin taxa (Fig. 3), as reported by Strange & Burr (1997). Paraphyly of *F. catenatus* with respect to the Mobile Basin taxa would be unusual indeed, given the morphological similarity amongst populations of *F. catenatus* and the fact that the two Mobile Basin taxa were not recognized as separate species until allozyme studies indicated a substantial genetic difference between specimens collected from the Tallapoosa and the Coosa and Alabama rivers (Rogers & Cashner, 1987; Cashner, Rogers & Grady, 1988). To investigate these preliminary results further, we used a multilocus approach

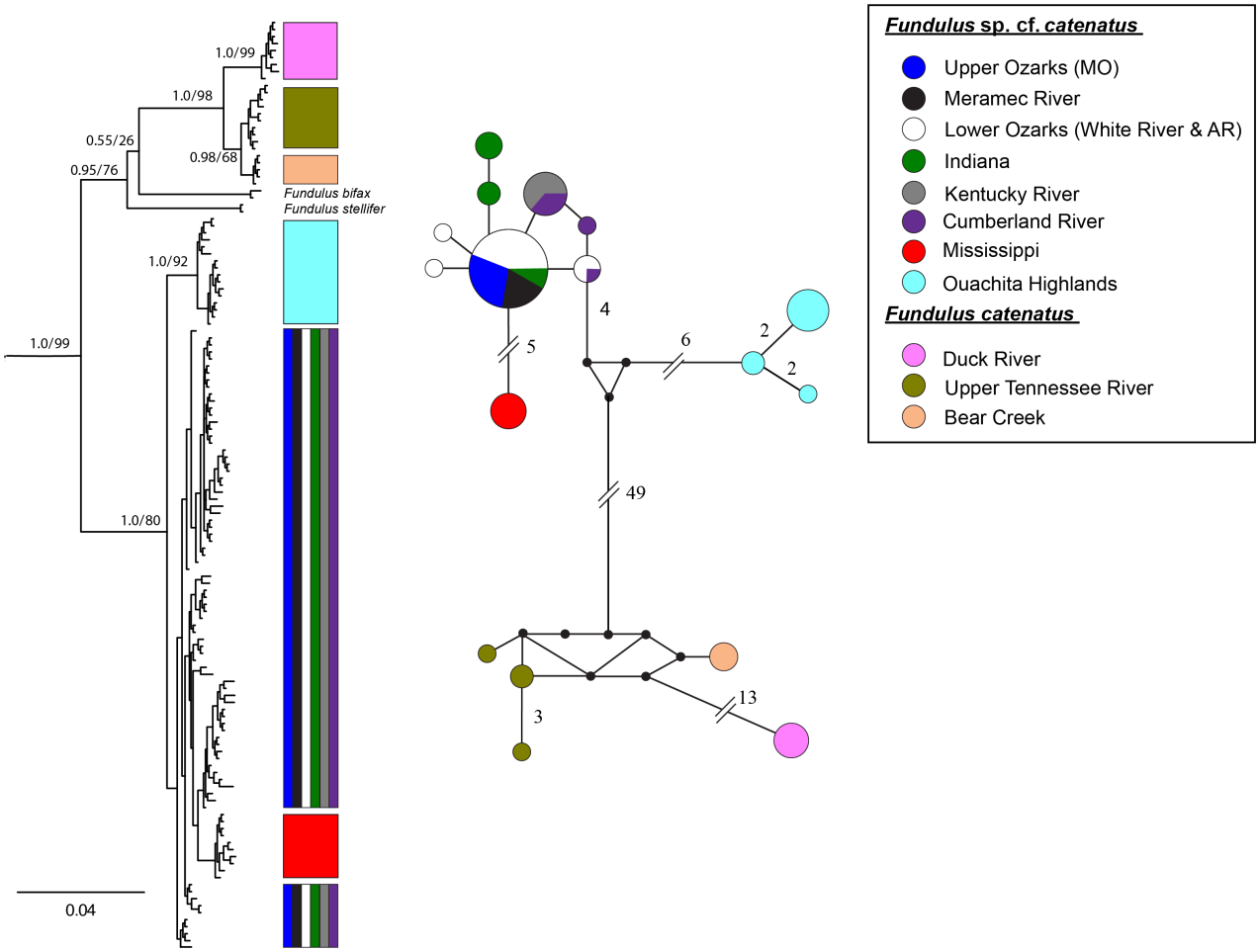

Figure 3. Phylogeny of unique haplotypes produced from Bayesian inference of Cytochrome *b* (cyt*b*) sequence data. Posterior probabilities followed by maximum likelihood bootstrap support at critical nodes. Scale bar is number of substitutions per site. *Cytb* haplotype network constructed by a median-joining method. The numbers annotated on branches of haplotype network represent the number of nucleotide base changes between haplotypes.

to resolve the phylogenetic relationships of the widely distributed North American studfishes, and to determine whether the two mitochondrial lineages observed were indicative of two species within *F. catenatus*. Our results found evidence for two species and, for clarity, we will refer to these as *F. catenatus*, restricted to the Buffalo and Tennessee rivers, and *F.* sp. cf. *catenatus,* occupying the rest of the range. We will refer to these two species combined as northern studfishes (*F. catenatus* + *F.* sp. cf. *catenatus*).

MATERIAL AND METHODS

Taxon sampling and DNA extraction, amplification, and sequencing

We obtained 24 samples of *F. catenatus* from six localities and 104 *F.* sp. cf. *catenatus* from 23 localities throughout their range and samples of *F. bifax* and *F. stellifer* from single localities (Fig. 2, Table 1). Samples of *Fundulus chrysotus* Günther (one individual), *Fundulus diaphanus* Lesueur (two), *Fundulus heteroclitus* Linnaeus (one), *Fundulus julisia* Williams & Etnier (two), *Fundulus rathbuni* Jordan & Meek (two), *Fundulus seminolis* Girard (two), and *Fundulus zebrinus* Jordan & Gilbert (two) were included as outgroups (Table 1).

Total genomic DNA was extracted from muscle or fin clips using Qiagen DNAeasy Blood and Tissue Kit (Qiagen, Valencia, CA) according to the manufacturer's suggested protocol. We used PCR to amplify one mtDNA locus, *cytb*, from all individuals and four nuclear DNA (nDNA) intron loci, syntaxin 5A (*stx5a*), nicalin (ncl1), ribosomal protein SA (*rpsa*), and ribosomal protein 3 (*rps3*) from a subset of samples (Table 1). All PCRs were performed in solution with total volume 12.5 μL, containing 1.5 μL

Table 1. Species, drainage, locality (from Fig. 1), river, latitude, longitude and number of individuals sampled for Cytochrome *b* and number of individuals subsampled for four nuclear DNA (nDNA) loci in this study.

Species	Drainage	Locality	River	Latitude (N)	Longitude (W)	Individuals (*cytb*)	Individuals (nDNA)
Fundulus sp. cf. *catenatus*	Homochitto & Amite Rivers (Mississippi)	1	East Fork Amite River	31°05′55″	90°43′11″	5	1
		2	Homochitto River	31°39′47″	90°43′14″	5	
	Ouachita River	3	Caddo River	34°23′10″	93°36′30″	5	1
		4	South Fork Ouachita River	34°33′29″	93°41′46″	5	
	Saline River	5	Ten Mile Creek	34°32′43″	92°45′15″	6	
	Arkansas River	6	Elk River	36°37′25″	94°35′25″	5	
		7	Elk River	36°33′37″	94°25′26″	5	1
	White River	8	Richland Creek	36°02′49″	93°58′19″	5	
		9	Kings River	36°23′39″	93°38′10″	5	1
		10	Strawberry River	36°05′56″	91°36′32″	4	
		11	Black River	37°25′00″	90°49′31″	3	1
		12	Current River	36°37′04″	90°50′22″	3	1
	Cuivre River	13	Big River	37°48′46″	90°46′20″	5	
	Meramec River	14	Huzzah Creek	37°56′52″	91°10′39″	5	
	Gasconade River	15	Big Piney River	37°13′11″	92°00′17″	10	1
	East Fork White River (Indiana)	16	Leatherwood Creek	39°34′56″	85°59′05″	5	1
		17	Sugar Creek	39°37′26″	85°56′46″	5	
	Green River	18	Trammel Fork	36°45′08″	86°17′15″	3	
		19	Falling Timber Creek	36°55′27″	85°48′19″	3	1
	Cumberland River	20	Hurricane Creek	36°35′58″	85°40′08″	3	
		21	Turnbull Creek	36°06′03″	87°07′35″	5	
		22	East Fork Stones River	35°56′30″	86°22′36″	3	1
		23	Otter Creek	36°42′47″	84°57′42″	2	
Fundulus catenatus	Lower Tennessee River	24	Bear Creek	34°38′01″	88°09′21″	5	2
	Buffalo River	25	Buffalo River	35°27′48″	87°32′07″	5	1
		26	Duck River	35°28′59″	86°27′47″	4	1
	Upper Tennessee River	27	Little River	35°47′07″	83°53′01″	4	1
		28	Little Pigeon River	35°52′12″	83°34′03″	4	1
		29	Clinch River	36°31′25″	83°09′20″	2	
Fundulus stellifer	Coosa River	30	Hachemedega Creek	32°50′42″	86°13′36″	2	2

Table 1. *Continued*

Species	Drainage	Locality	River	Latitude (N)	Longitude (W)	Individuals (*cytb*)	Individuals (nDNA)
Fundulus bifax	Tallapoosa River	31	Josie Leg Creek	32°59′08″	85°51′36″	2	2
Fundulus chrysotus		Outgroup	Saline River	34°19′10″	92°35′13″	1	
Fundulus diaphanus		Outgroup	Mille Lacs Lake	46°14′47″	93°39′59″	2	1
Fundulus heteroclitus		Outgroup	GenBank (FJ445402)			1	
Fundulus julisia		Outgroup	Conservation Fisheries, Inc. (Charles Creek)			2	1
Fundulus rathbuni		Outgroup	Country Line Creek	36°21′11″	79°17′57″	2	1
Fundulus seminolis		Outgroup	Lake Okeechobee	27°11′51″	80°49′49″	1	
		Outgroup	Lake George	29°12′05″	81°34′40″	1	
Fundulus zebrinus		Outgroup	Willow Creek	38°56′24″	101°57′48″	2	

template DNA, 2.75 μL water, 6.25 μL GoTaq Green Master Mix (Promega, Madison, WI), 1.0 μL forward primer, and 1.0 μL reverse primer. *Cytb* was amplified using the GLU (5′-GACTTGAAGAACCA CCGTTG-3′) and THR (5′-TCCGACATTCGGTTTAG AAG-3′) primers described in Near, Porterfield & Page (2000) and using the following thermal conditions: denaturation at 94 °C for 30 s, annealing at 50 °C for 30 s, and extension at 72 °C for 90 s, repeated for 25 cycles.

The four nDNA intron loci were amplified using the following primers: *stx5a* – 10F (5′-GGAGGAG ACKGACTGGAAGT-3′) and 10R (5′-GCAGAACATY GARAGCACMA-3′), ncl1 – 428 bp F8 (5′-SGCCAGG TTGATYTTCTTRT-3′) and R8 (5′-CCAGTCTGCTSC AGGACAAY-3′), *rpsa* – 3F (5′-ATTGTTGCCATYG ARAAYCC-3′) and 3R (5′-GCWGCCTGRATCTGA TTGGT-3′), and *rps3* – F4 (5′-CTACAAGCTGCTSGG AGGMC-3′) and R4 (5′-TAGTTSACKGGGTCTCC RCT-3′) (Halas & Simons, 2014, D. Halas pers. comm.). All introns were amplified with an initial denaturation at 95 °C for 4 min, followed by denaturation at 95 °C for 40 s, varying annealing temperatures for 40 s, and 7 °C for 90 s for 25 cycles. Annealing temperature for ncl1, *stx5a*, *rpsa*, and *rps3* was 50, 51, 52, and 56 °C, respectively. Owing to initial difficulty phasing *rpsa* and *rps3* sequences of some individuals, internal primers were developed using PRIMER3 (http://frodo.wi.mit.edu/primer3/). Additional *rpsa* sequences were amplified using the intron thermal profile with the primer pairs F101 (5′-GTAAACGGATCGGGGTTTCT-3′) and R800 (5′-AAGGCCCTTTTCACTTTTCA-3′), and F146 (5′-TG ACTGGGGTATGAGAAGCTC-3′) and R728 (5′-CA CGCTTTCTAACCTCCCTTT-3′) and an annealing temperature of 50 and 54 °C, respectively.

PCR products were purified using Exonuclease 1 and shrimp alkaline phosphatase (USB Corporation, Cleveland, OH) at the manufacturer's suggested thermal profiles. Automated Sanger sequencing of purified PCR products was performed using ABI Prism BigDye Terminator v. 3.1 chemistry (Applied Biosystems, Foster City, CA) at the Biomedical Genomics Center DNA Sequencing and Analysis Facility at the University of Minnesota.

Editing and aligning sequences

Complementary heavy and light strands were aligned into contiguous sequences (contigs) and edited in GENEIOUS v. 6.1.6 (www.geneious.com; Biomatters Ltd., Auckland, New Zealand). Length heterozygotes found in nuclear introns were phased by eye with help from CHAMPURU v. 1.0 (Flot, 2007, available online at http://www.mnhn.fr/jfflot/champuru/). Consensus sequences of contigs were

aligned using the MUSCLE (Edgar, 2004) clustering algorithm as implemented in GENEIOUS v. 6.1.6. Each alignment was trimmed to make sequences near uniform in length. In intron alignments, large indels found only in outgroups were removed. This reduced the length of *rps3* and *stx5a*. Nuclear introns were tested for recombination using the phi test (Bruen, Philippe & Bryant, 2006) as implemented in SplitsTree v. 4.13.1 (Huson & Bryant, 2006). For some analyses, identical *cytb* haplotypes were removed using ElimDupes (http://hcv.lanl.gov/content/sequence/ELIMDUPES/elimdupes.html).

Phylogenetic analysis

Two data sets were created for *cytb:* all individuals and unique haplotypes. The best-fitting partitioning scheme and nucleotide substitution models were determined using PartitionFinder v. 1.01 (Lanfear *et al.*, 2012) based on Bayesian information criterion scores. Bayesian analysis of each *cytb* data set was conducted using MrBayes v. 3.2.1 (Ronquist *et al.*, 2012) on the CIPRES Science Gateway portal (Miller, Pfeiffer & Schwartz, 2010). The Metropolis coupled Markov chain Monte Carlo (MCMCMC) command was used for two simultaneous runs with four chains (three heated chains, one cold) per 15 000 000 generations, sampling every 1000. Log files were checked in the program TRACER v. 1.5 (http://beast.bio.ed.ac.uk/Tracer) to assess convergence of runs and burn-in was set to remove the first 20% of sampled trees. We performed maximum likelihood (ML) analyses in the program GARLI v. 2.0 (Zwickl, 2006) on the CIPRES Science Gateway portal (Miller *et al.*, 2010). The gene tree with the best likelihood score was selected from five search replicates. The nodes of the best ML tree found by the aforementioned five search replicates were annotated with the proportion of nodes found by 1000 bootstrap replicates using SumTrees v. 3.3.1 in the DendroPy v. 3.11 package (Sukumaran & Holder, 2010).

To account for genetic diversity observed in the *cytb* gene tree, we subsampled six *F. catenatus*, ten *Fundulus* sp. cf. *catenatus,* two *F. bifax* and *F. stellifer*, and one *F. julisia*, *F. diaphanus*, and *F. rathbuni* for the nuclear intron data set. Phylogenetic analysis of each of the four nuclear intron data sets followed the same methods of partitioning scheme and nucleotide substitution model selection, Bayesian, and ML analyses.

Species tree analysis

Species tree analysis of the subsampled individuals (six *Fundulus catenatus,* ten *F.* sp. cf. *catenatus*, two *F. bifax* and *F. stellifer*, and one *F. julisia*, *F. diaphanus*, and *F. rathbuni*) was conducted using *BEAST v. 1.7.5 (Heled & Drummond, 2010). Species were designated based on the hypothesis of relationship suggested by the *cytb* gene tree. The following methods were used for an all loci (*cytb* + nuclear introns) and just nuclear intron data set. An appropriate clock model was determined by performing molecular clock likelihood ratio tests for each locus in PAUP* v. 4.0b10 (Swofford, 2003). A Yule process speciation prior was used for branching rates. We applied similar partitioning schemes and nucleotide substitution models to analyses of individual genes. Ten independent runs of 50 000 000 generations each were conducted, sampling every 1000 generations. The MCMCMC log files were analysed in TRACER v. 1.5 (http://beast.bio.ed.ac.uk/Tracer) to assess convergence of the runs, ensure proper mixing, and determine an appropriate burn-in (first 10% of sampled trees). LogCombiner v. 1.7.4 (http://beast.bio.ed.ac.uk/LogCombiner) was used to remove burn-in and combine files; the resulting 10 000 trees were used to produce a maximum clade credibility tree using TreeAnnotator v. 1.7.4 (http://beast.bio.ed.ac.uk/TreeAnnotator).

Species delimitation using nDNA

To assess the deep split recovered in both the *cytb* gene tree and species tree between *Fundulus catenatus* and *F.* sp. cf. *catenatus*, we used the program BAYESIAN PHYLOGENETICS AND PHYLOGEOGRAPHY (BP&P v. 2.2: Yang & Rannala, 2010) to compare a one-species model (*F. catenatus* combined with *F.* sp. cf. *catenatus*) and a two-species model (*F. catenatus* and *F.* sp. cf. *catenatus*). In BP&P, we used the reversible-jump Markov chain Monte Carlo method (Rannala & Yang, 2013) to delimit species and assess cryptic diversity. BP&P requires a guide tree, species group membership definitions, and sequence alignments. The guide tree and species group memberships were assigned according to our hypothesis of species limits based on the *cytb* gene tree (Fig. 2). The phased nuclear introns of six *F. catenatus* and ten *F.* sp. cf. *catenatus* individuals were provided as the sequence alignments. To evaluate the influence of some priors and settings we ran multiple runs for 500 000 generations, sampling every five generations, and used a burn-in of 50 000 generations. We considered three different combinations of prior distributions of two parameters known to influence the posterior probability for models, ancestral population size (Θ) and root age (τ) (Yang & Rannala, 2010): large ancestral population size with deep divergence amongst species, small ancestral population size with shallow divergence amongst species, and large ancestral population size with small

divergence amongst species (Leaché & Fujita, 2010). As suggested by the user's manual we set the mutation rate parameter, 'locusrate', at 15 to account for similar rates amongst noncoding loci and ran analysis utilizing both provided *reversible-jump Markov chain Monte Carlo algorithms* to ensure similar results. A posterior probability ≥ 0.95 at the splitting event of *F. catenatus* and *F.* sp. cf. *catenatus* was considered strong evidence of two species (following Leaché & Fujita, 2010).

To further explore evidence for multiple species, we utilized a species delimitation method using Bayes factors described by Grummer, Bryson & Reeder (2014). This method compares marginal likelihood scores, estimated with stepping-stone sampling (Xie *et al.*, 2011), and path sampling (Lartillot & Philippe, 2006), using Bayes factors. We estimated species trees for both species delimitation models using the methods presented in the Species Tree Analysis section. The sole difference was that the initial species tree inference was run for 75 000 000 generations, sampling every 1000 generations. Following *BEAST analysis, the initial 20% of sampling was removed as burn-in, and both path sampling and stepping-stone sampling were executed for a chain length of 7 500 000 generations for 300 paths (totalling 225 000 000 generations). This was carried out five times to provide evidence of consistent results. Resulting marginal likelihood scores were compared as 2Ln (Bayes factors) (where Bayes factor = marginal likelihood score of one-species model – marginal likelihood score of a two-species model). Kass & Raftery (1995) suggested considering 2Ln (Bayes factors) scores of 0–2, 2–6, 6–10, and > 10 as 'not worth more than a bare mention', 'positive' support, 'strong' support, and 'very strong' support, respectively.

Population statistics and median-joining network

The following summary statistics were assessed using DnaSP v. 5.10.1 (Librado & Rozas, 2009): singletons, polymorphic sites, parsimony informative sites, number of haplotypes, haplotype diversity, and nucleotide diversity.

A *cytb* haplotype network was constructed by the median-joining method (Bandelt, Forster & Röhl, 1999) in the program NETWORK v. 4.6.1.2 (fluxus-engineering.com). The data set contained all individuals sampled and was trimmed to the shortest sequences (934 bp) because large amounts of missing data in a sequence led to a large number of unverifiable 'unique' haplotypes. All default settings were used with the one exception of the switching parameters frequency > 1 criterion being set to active.

Neutrality tests

We explored possible population expansion using a coalescent-based approach with Fu's F_S (Fu, 1997) and Ramos-Onsins and Rozas' R_2 (Ramos-Onsins & Rozas, 2002). We used both methods as we have varying sample sizes, and tests run by Ramos-Onsins & Rozas (2002) found R_2 to function better for small sample sizes, whereas F_S behaved best for large sample sizes. The following populations were selected based on median-joining network: *F.* sp. cf. *catenatus* (excluding Mississippi and Ouachita), *F.* sp. cf. *catenatus* (Ouachita), and *F. catenatus*. The southern Mississippi population of *F.* sp. cf. *catenatus* was excluded from this analysis owing to limited sampling (ten individuals). F_S and R_2 were calculated and coalescent simulations were run given segregating sites for 10 000 replicates in the program DnaSP v. 5.10.1. Significantly negative values of F_S (P-value < 0.02) and significant R_2 (<0.05) suggest an excess of rare haplotypes, indicative of non-neutral processes such as recent demographic expansion or genetic hitchhiking (Fu, 1997; Ramos-Onsins & Rozas, 2002).

RESULTS

Sequencing and alignments

We successfully sequenced *cytb* for 24 individuals of *F. catenatus* and 104 individuals of *F.* sp. cf. *catenatus* and subsampled six *F. catenatus* and ten *F.* sp. cf. *catenatus* to represent major clades for all four nuclear introns**.** The longest sequences in each gene alignment were 1085 bp (*cytb*), 478 bp (*stx5a*), 428 bp (*ncl1*), 999 bp (*rpsa*), 1536 bp (*rps3*) (Supporting Information Table S1). The *cytb* data set was reduced from 128 to 59 individuals after removing duplicate haplotypes. *Rps3* and *stx5a* contained long indel regions, which were found only in the outgroup species *F julisia* and *F. rathbuni*, and subsequently removed from the alignments prior to all analyses.

Model selection

The best-fitting partitioning schemes and models of nucleotide substitution selected for use in Bayesian and maximum likelihood gene tree analyses were as follows: cyt*b* by codon position (F81, GTR + Γ, K80 + Γ), *ncl1* (JC + Γ), *rps3* (HKY + Γ), *rpsa* (HKY + Γ), and *stx5a* (K80 + Γ). The models of nucleotide substitution were adjusted as follows to fit options available in the program *BEAST: cyt*b* (all positions GTR + Γ) and all nuclear introns (HKY + Γ).

Phi tests detected no evidence for recombination in any of the nuclear genes at a significance level of 0.05. The likelihood ratio test rejected the molecular

clock for *cytb* and *rps3*. As such, in the species tree analyses *cytb* and *rps3* were assigned a relaxed clock uncorrelated exponential model and *ncl1*, *rpsa*, and *stx5a* were assigned a strict molecular clock model.

Mitochondrial and nuclear gene trees

The monophyly of the *F. catenatus* species group (*F. catenatus*, *F.* sp. cf. *catenatus, F. bifax,* and *F. stellifer*) was strongly supported by the cyt*b* gene tree (posterior probability (pp) = 1.0, ML = 99; Fig. 3). *Fundulus catenatus*, *F. bifax,* and *F. stellifer* form a well-supported clade (pp = 0.95, ML = 76) sister to *F.* sp. cf. *catenatus* (pp = 1.0, ML = 80; Fig. 1). Within *F. catenatus* there are three well-supported clades formed by individuals from Bear Creek in far north-east Mississippi (pp = 0.99, ML = 76), Upper Tennessee River (upstream of Guntersville, AL; pp = 0.99, ML = 68), and Buffalo River (including tributaries such as the Duck River; pp = 0.99, ML = 99) (Fig. 3). *Fundulus* sp. cf. *catenatus* had less geographical structure, as only the isolated Mississippi (pp = 1.0, ML = 97) and Ouachita individuals (pp = 1.0, ML = 92) formed well-supported clades.

The gene trees for all nuclear introns had some geographical structure (Fig. S1). *Fundulus catenatus* and *F.* sp. cf. *catenatus* were often recovered as monophyletic, but not reciprocally so, and the relationship between these two clades was not fully resolved for any nuclear locus. All gene trees, except *stx5a*, provided strong support for the monophyly of the *F. catenatus* species group (*F. catenatus*, *F.* sp. cf. *catenatus*, *F. bifax*, and *F. stellifer*).

Species tree analysis

Species tree analysis of all loci in *BEAST provided strong support for the *F. catenatus* species group (pp = 0.98) and a clade with *F. catenatus*, sister to *F.* sp. cf. *catenatus* (pp = 0.99; Fig. 4A). Species tree analysis of just nuclear loci in *BEAST provided the same topology with reduced support for each clade. In this analysis the sister relationship between *F. bifax* and *F. stellifer* is not strongly supported (pp = 0.72, 0.68 respectively; Fig. 4B).

Species delimitation

There is substantial evidence for two *Fundulus* species currently recognized as *F. catenatus*: *F. catenatus* in the Tennessee River and its tributaries, and *F.* sp. cf. *catenatus,* occupying the rest of the range. All posterior probability distributions for all BP&P runs delimited *F. catenatus* and *F.* sp. cf. *catenatus* as separate species (pp = 1.0). These results were replicated by all iterations tested on prior distributions of two parameters known to influence the posterior probability for models, ancestral population size (Θ) and root age (τ). Additionally, most fine-tune MCMCMC move acceptance proportions were optimal, 0.30, and all were in the interval (0.15, 0.7) suggested by the BP&P manual.

A two-species model is strongly supported over a one-species model based on hypothesis testing using BayesFactors. All five independent path sampling and stepping-stone procedures sampling executed in *BEAST resulted in very similar marginal likelihood scores and ultimately Bayes factors greater than 10, all between the narrow range of 92–98.

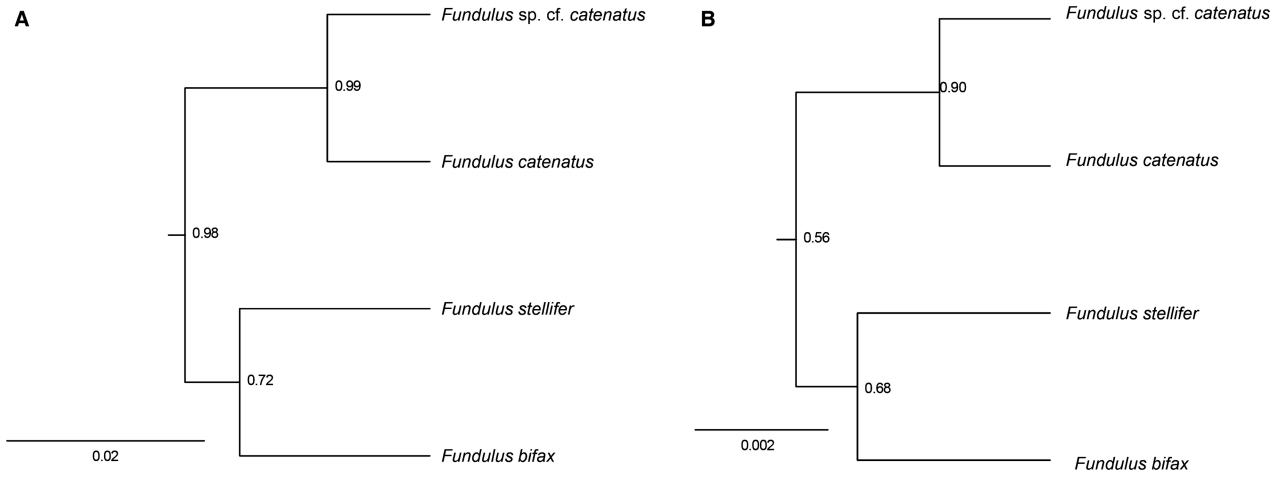

Figure 4. Phylogenies from species tree analysis conducted in *BEAST. Posterior probabilities listed at all nodes. A, all loci [cytochrome *b* + nuclear DNA (nDNA)]; B, nDNA. Scale bar represents substitutions per site.

Population statistics and median-joining network

The median-joining haplotype network for cyt*b* separated *F.* sp. cf. *catenatus* from *F. catenatus* by at least 49 nucleotide substitutions (Fig. 3). Within each of these two large clusters is some geographical structure. For example, *F.* sp. cf. *catenatus* is divided into two large clusters; one from the Ouachitas separated by at least six nucleotide substitutions from the other cluster with several widespread haplotypes and one restricted haplotype (Mississippi population). Within *F. catenatus* there appear to be three distinct populations including one separated by at least 13 nucleotide substitutions from its nearest haplotype.

Neutrality tests

Fundulus sp. cf. *catenatus* (minus Mississippi and Ouachita) was the only group to exhibit significantly negative Fu's F_S ($F_S = -21.62$, *P*-value 0.0001) and significant R_2 ($R_2 = 0.03$, *P*-value = 0.0003; Table 2). *Fundulus catenatus* from Ouachita and *F. catenatus* (Tennessee River) both had nonsignificant results for Fu's F_S and R_2 (Table 2).

DISCUSSION

Phylogenetic analysis of cyt*b* sequence data indicates that *F. catenatus sensu lato* is paraphyletic, corroborating Strange & Burr (1997). The observed groups largely conflict with 'populations' based on morphological variation reported by Thomerson (1969), but clearly indicate the presence of four distinct studfish lineages: *F. bifax* and *F. stellifer* in the Mobile Basin, *F. catenatus* in the Tennessee River and its tributaries, and an undescribed species, *F.* sp. cf. *catenatus,* present across the rest of the range (Fig. 3). The short branches at the base of the cyt*b* tree may indicate that the four taxa rapidly diverged. In this situation, one would expect differences between gene trees and species trees, although mitochondrial and nuclear gene tree discord is generally rare (Bowen *et al.*, 2014).

Species tree analysis, based on nuclear loci (nDNA) and all loci (cyt*b* + nDNA), supports a monophyletic *F. catenatus* species group (Fig. 4), consistent with the analyses of mitochondrial sequences. The support for monophyly of the Mobile Basin taxa was low (pp = 0.72), again suggesting that speciation within the *F. catenatus* group occurred over a short time period (Fig. 4A). However, unlike the cyt*b* tree, *F. catenatus* plus *F.* sp. cf. *catenatus* is monophyletic, sister to a monophyletic group containing the Mobile Basin taxa, *F. bifax* and *F. stellifer*. This hypothesis is consistent with phylogenetic patterns

Table 2. Population statistics including number of sequences per site, number of singletons, number of polymorphic sites, parsimony informative sites, haplotype diversity, nucleotide diversity, Fu's F_S, and R_2 based on cytochrome *b* sequences (1085 bp).

Group	*N*	Singletons	Polymorphic sites (segregating)	Parsimony informative sites	Number of haplotypes	Haplotype diversity	Nucleotide diversity	Fu's F_s (*P*-value)	R_2 (*P*-value)
Fundulus sp. cf. *catenatus* (minus Mississippi and Ouachita)	79	23	35	12	27	0.795	0.00229	−21.621 ($P < 0.001$*)	0.0305 ($P < 0.001$*)
Fundulus sp. cf. *catenatus* (Ouachita)	16	2	6	4	4	0.592	0.00184	1.015 ($P = 0.753$)	0.1465 ($P = 0.527$)
Fundulus catenatus (Tennessee River)	24	5	30	25	12	0.909	0.01061	1.158 ($P = 0.717$)	0.1761 ($P = 0.936$)

*$P < 0.05$.

in other fish taxa (Berendzen, Simons & Wood, 2003; Near & Keck, 2005; Roe, Mayden & Harris, 2008) and is reflected in hypotheses of drainage relationships in synthetic biogeographical analyses (Mayden, 1988; Hoagstrom, Ung & Taylor, 2014). Near & Keck (2005) and Roe *et al.* (2008) suggested a vicariance event separating what is now the Upper Tennessee River from the Mobile Basin, in the middle to late Miocene, consistent with the formation of a drainage system similar to the modern Tennessee River (Galloway, Whiteaker & Ganey-Curry, 2011; Hoagstrom *et al.*, 2014).

The species delimitation analyses, based on nuclear loci, corroborate the hypothesis of cryptic diversity. We argue that the northern studfishes comprise at least two cryptic taxa, one occupying the Tennessee and Buffalo rivers, and a second occupying the rest of the range. The type specimen of *F. catenatus* was collected in Florence, Alabama (Storer, 1846); thus, the Tennessee and Buffalo rivers contain the nominal species. Studies based on morphological and molecular data have demonstrated other cryptic species in the Eastern Highlands and recent work has demonstrated that phylogeographical patterns amongst taxa exhibit similarities but rarely congruence (Strange & Burr, 1997; Harrington & Near, 2012; Halas & Simons, 2014), indicating that the rivers of the Eastern Highlands have a long and complex geographical history.

Analyses of the cyt*b* data indicate that *F. catenatus* has substantial population structure with evidence for at least three distinct populations, Buffalo River, Upper Tennessee River, and Bear Creek, suggesting little or no gene exchange between them. It is unclear why *F. catenatus* would exhibit so much genetic structure across a relatively small geographical area, particularly as many of the other taxa inhabiting the Tennessee and Cumberland rivers show genetic subdivision within each of those systems (Hollingsworth & Near, 2009; Keck & Near, 2010; Halas & Simons, 2014). This is in stark contrast to *F.* sp. cf. *catenatus*, which exhibits very little genetic structure across a much larger range. This may reflect an interaction between dispersal behaviour in *F. catenatus* and the physiography of the Tennessee and Buffalo rivers. It may also reflect species-specific dispersal behaviours between the two species. Differences in dispersal between closely related species was observed in the *Etheostoma zonale* species group, with *E. zonale* replacing a closely related cryptic species via introgression as *E. zonale* expands its range upstream (Halas & Simons, 2014).

Fundulus sp. cf. *catenatus* occupies a much greater geographical area (Figs 2, 3) than *F. catenatus*. Three genetically distinct clusters are evident: Ouachita Highlands, Mississippi, and remaining samples from the rest of the range. The Ouachita Highlands haplotypes form a monophyletic group, sister to all remaining haplotypes of *F.* sp. cf. *catenatus*. The Ouachita and Red rivers drain the Ouachita Highlands and contain several endemic highland fish species, or distinct populations (Mayden, 1985), indicative of a long period of isolation. Several clades exhibit patterns similar to *F.* sp. cf. *catenatus,* where the Ouachita and Red river taxa are sister to a larger and more widely distributed clade (Berendzen *et al.*, 2008; Blum *et al.*, 2008; Egge & Simons, 2011). Other taxa with populations in the Ouachita Highlands do not indicate differentiation between the Ouachita Highlands and other areas (Halas & Simons, 2014), suggesting that the Ouachita Highlands taxa are not the same age and that there have been multiple dispersals into this area.

Fundulus sp. cf. *catenatus* isolated in southern Mississippi are also monophyletic but are nested within the non-Ouachita *F.* sp. cf. *catenatus* clade (Fig. 3). Other relict highland taxa are present in and near the Homochitto-Amite basins, including *Chrosomus erythrogaster* (S. Martin, pers. comm.), *Hypentelium nigricans* (Berendzen *et al.*, 2003), and *Etheostoma caeruleum* (Ray, Wood & Simons, 2006). Unlike *F.* sp. cf. *catenatus*, these taxa are nested within a widespread Mississippi River clade rather than forming a distinct monophyletic group.

The remainder of the range of *F.* sp. cf. *catenatus*, extending from eastern Kentucky, west to Oklahoma and from Arkansas north to Indiana (Fig 2), is characterized by the presence of widespread haplotypes with no evidence of haplotypes restricted to particular geographical locations (Fig. 3). The haplotype distributions and significantly negative values for F_S and R_2 suggest that *F.* sp. cf. *catenatus* (minus the Mississippi and Ouachita samples) has undergone recent population expansion accompanied by dispersal (Table 2). This dispersal was probably followed by at least one contraction of the range, leaving individuals isolated in small tributaries to the East Fork White River, Indiana. There are many examples of northern expansion of the ranges of highland fishes into the upper Mississippi and upper Ohio river systems following the Pleistocene glaciations (Near, Page & Mayden, 2001; Simons, 2004; Echelle *et al.*, 2014; Halas & Simons, 2014). *Fundulus* sp. cf. *catenatus* differs from these in that it has not dispersed extensively into previously glaciated areas. It is possible that the population expansion of *F.* sp. cf. *catenatus* is a relatively recent phenomenon, occurring after northward dispersal corridors were no longer accessible. There is some evidence for ongoing expansion of the range of this taxon via stream capture and other mechanisms. Thomerson (1969) mentioned a number of recent locality records that suggest

F. sp. cf. *catenatus* is expanding its range in the Ozark Highlands via unknown means, and Branson & Batch (1971) hypothesized recent range expansion into the Kentucky River system via stream captures from the Salt River.

This study was designed to investigate the observation of paraphyly of *F. catenatus* and *F.* sp. cf. *catenatus* with respect to Mobile Basin taxa and thus did not address other potentially cryptic taxa in this group. Within *F. catenatus*, the large number of steps in the haplotype network between samples from the Duck River, Upper Tennessee River and Bear Creek (Fig. 3) suggests an extensive period of isolation. This area is rife with endemic taxa with restricted ranges (Harrington & Near, 2012; Keck & Near, 2013) and other cryptic taxa have been identified from the Upper Tennessee River (Halas & Simons, 2014). It is possible that *F.* sp. cf. *catenatus* also contains additional cryptic diversity, notably the Ouachita Highlands population. The cyt*b* tree and haplotype network demonstrate that this group is monophyletic, sister to the rest of *F.* sp. cf. *catenatus.* The Ouachita Highlands individuals were identified as morphologically distinct by Thomerson (1969), and again, this area has a number of endemic taxa. Furthermore, our sampling did not include the upper Red River of Arkansas and Oklahoma. In order to further explore species limits in Ouachita Highlands *F.* sp. cf. *catenatus* and the three distinct populations of *F. catenatus* more individuals and molecular loci need to be sampled.

CONCLUSION

We used a multilocus approach to identify cryptic diversity and resolve the phylogenetic relationships of the widely distributed North American studfishes. We discovered one species restricted to the Tennessee River and its tributaries, and a second, widespread species that has undergone a recent population expansion. It is possible that both of these species contain additional cryptic taxa, but a critical assessment of this is beyond the reach of our data. We were able to corroborate our hypothesis of cryptic diversity, which derived from analysis of mitochondrial data and demonstrated that the phylogenetic relationships identified using only mitochondrial data were misleading, probably because the events leading to isolation and speciation occurred over a relatively short period of time.

ACKNOWLEDGEMENTS

We thank Andrew P. Kinzinger, Bernard R. Kuhajda, John F. Switzer, Robert M. Wood, and Malcolm Johns for their help collecting specimens used in this study. The States of Alabama, Arkansas, Indiana, Kentucky, Mississippi, Missouri, and Tennessee provided scientific collecting permits. Brett C. Nagle, M. Vincent Hirt, and Robert M. Wood provided thoughtful comments on the manuscript. Dominik Halas provided invaluable help phasing alleles. We thank Joseph R. Tomelleri for use of his illustrations. Funding was provided by the Bell Museum of Natural History, the Department of Fisheries, Wildlife, and Conservation Biology, and the Minnesota Agricultural Experiment Station.

REFERENCES

Bandelt H-J, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Berendzen PB, Simons AM, Wood RM. 2003. Phylogeography of the northernhogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. *Journal of Biogeography* **30:** 1139–1152.

Berendzen PB, Simons AM, Wood RM, Dowling TE, Secor CL. 2008. Recovering cryptic diversity and ancient drainage patterns in eastern North America: historical biogeography of the *Notropis rubellus* species group (Teleostei: Cypriniformes). *Molecular Phylogenetics and Evolution* **46:** 721–737.

Blum MJ, Neely DA, Harris PM, Mayden RL. 2008. Molecular systematics of the cyprinid genus *Campostoma* (Actinopterygii: Cypriniformes): disassociation between morphological and mitochondrial differentiation. *Copeia* **2008:** 360–369.

Bowen BW, Shanker K, Yasuda N, Malay MCD, von der Heyden S, Paulay G, Rocha LA, Selkoe KA, Barber PH, Williams ST, Lessios HA, Crandall ED, Bernardi G, Meyer CP, Carpenter KE, Toonen RJ. 2014. Phylogeography unplugged: comparative surveys in the genomic era. *Bulletin of Marine Science* **90:** 13–46.

Branson BA, Batch DL. 1971. Stream capture in Kentucky indicated by distributional records of *Fundulus catenatus* and *Etheostoma spectabile*. *American Midland Naturalist* **86:** 496–500.

Bruen TC, Philippe H, Bryant D. 2006. A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172:** 2665–2681.

Carstens BC, Pelletier TA, Reid NM, Satler JD. 2013. How to fail at species delimitation. *Molecular Ecology* **22:** 4369–4383.

Cashner RC, Rogers JS, Grady JM. 1988. *Fundulus bifax*, a new species of the subgenus *Xenisma* from the Tallapoosa and Coosa river systems of Alabama and Georgia. *Copeia* **1988:** 674–683.

Echelle AA, Schwemm MR, Lang NJ, Nagle BC, Simons AM, Unmack PJ, Fisher WL, Hoagstrom CW. 2014. Molecular systematics and historical biogeography of the *Nocomis biguttatus* species group (Teleostei: Cyprinidae): nuclear and mitochondrial introgression and a cryptic Ozark species. *Molecular Phylogenetics and Evolution* **81:** 109–119.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32:** 1792–1797.

Egge JJD, Simons AM. 2006. The challenge of truly cryptic diversity: diagnosis and description of a new madtom catfish (Ictaluridae: Noturus). *Zoologica Scripta* **35:** 581–595.

Egge JJD, Simons AM. 2011. Evolution of venom delivery structures in madtom catfishes (Siluriformes: Ictaluridae). *Biological Journal of the Linnean Society* **102:** 115–129.

Flot JF. 2007. Champuru 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths. *Molecular Ecology Notes* **7:** 974–977.

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

Galloway WE, Whiteaker TL, Ganey-Curry P. 2011. History of Cenozoic North American drainage basin evolution, sediment yield, and accumulation in the Gulf of Mexico Basin. *Geosphere* **7:** 938–973.

Ghedotti MJ, Davis MP. 2013. Phylogeny, classification, and evolution of salinity tolerance of the North American topminnows and killifishes, family Fundulidae (Teleostei: Cyprinodontiformes). *Fieldiana Life and Earth Sciences* **7:** 1–65.

Ghedotti MJ, Simons AM, Davis MP. 2004. Morphology and phylogeny of the studfish clade, subgenus Xenisma (Teleostei: Cyprinodontiformes). *Copeia* **2004:** 53–61.

Grady JM, Cashner RC, Rogers JS. 1990. Evolutionary and biogeographic relationships of *Fundulus catenatus* (Fundulidae). *Copeia* **1990:** 315–323.

Grummer JA, Bryson RW, Reeder TW. 2014. Species delimitation using Bayes factors: simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). *Systematic Biology* **63:** 119–133.

Halas DS, Simons AM. 2014. Cryptic speciation reversal in the *Etheostoma zonale* (Teleostei: Percidae) species group, with an examination of the effect of recombination and introgression on species tree inference. *Molecular Phylogenetics and Evolution* **70:** 13–28.

Harrington RC, Near TJ. 2012. Phylogenetic and coalescent strategies of species delimitation in snubnose darters (Percidae: *Etheostoma*). *Systematic Biology* **61:** 63–79.

Heled J, Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27:** 570–580.

Hoagstrom CW, Ung V, Taylor K. 2014. Miocene rivers and taxon cycles clarify the comparative biogeography of North American highland fishes. *Journal of Biogeography* **41:** 644–658.

Hollingsworth PR, Near TJ. 2009. Temporal patterns of diversification and microendemism in Eastern Highland endemic barcheek darters (Percidae: Etheostomatinae). *Evolution* **63:** 228–243.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23:** 254–267.

Kass RE, Raftery AE. 1995. Bayes factors. *Journal of the American Statistical Association* **90:** 773–795.

Keck BP, Near TJ. 2010. Geographic and temporal aspects of mitochondrial replacement in *Nothonotus* darters (Teleostei: Percidae: Etheostomatinae). *Evolution* **64:** 1410–1428.

Keck BP, Near TJ. 2013. A new species of *Nothonotus* darter (Teleostei: Percidae) from the Caney Fork in Tennessee, USA. *Bulletin of the Peabody Museum of Natural History* **54:** 3–21.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29:** 1695–1701.

Lartillot N, Philippe H. 2006. Computing Bayes factors using thermodynamic integration. *Systematic Biology* **55:** 195–207.

Leaché AD, Fujita MK. 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B* **277:** 3071–3077.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25:** 1451–1452.

Mayden RL. 1985. Biogeography of Ouachita Highland fishes. *The Southwestern Naturalist* **30:** 195–211.

Mayden RL. 1988. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. *Systematic Zoology* **37:** 329–355.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE), 14 November 2010, New Orleans, LA: 1–8.

Myer GS. 1964. A brief sketch of the history of ichthyology in America to the year 1850. *Copeia* **1964:** 33–41.

Near TJ, Keck BP. 2005. Dispersal, vicariance, and timing of diversification in *Nothonotus* darters. *Molecular Ecology* **14:** 3485–3496.

Near TJ, Porterfield JC, Page LM. 2000. Evolution of cyt*b* and the molecular systematics of *Ammocrypta* (Percidae: Etheostomatinae). *Copeia* **2000:** 701–711.

Near TJ, Page LM, Mayden RL. 2001. Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. *Molecular Ecology* **10:** 2235–2240.

Niemiller ML, Near TJ, Fitzpatrick BM. 2012. Delimiting species using multilocus data: diagnosing cryptic diversity in the Southern cavefish, *Typhlichthys subterraneus* (Teleostei: Amblyopsidae). *Evolution* **66:** 846–866.

Ramos-Onsins SE, Rozas J. 2002. Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution* **19:** 2092–2100.

Rannala B, Yang Z. 2013. Improved reversible jump algorithms for Bayesian species delimitation. *Genetics* **194:** 245–253.

Ray JM, Wood RM, Simons AM. 2006. Phylogeography and post-glacial colonization patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). *Journal of Biogeography* **33:** 1550–1558.

Roe KJ, Mayden RL, Harris PM. 2008. Systematics and zoogeography of the rock basses (Centrarchidae: Ambloplites). *Copeia* **2008:** 858–867.

Rogers JS, Cashner RC. 1987. Genetic variation, divergence, and relationships in the subgenus *Xenisma* of the genus *Fundulus*. In: Matthews WJ, Heins DC, eds. *Community and evolutionary ecology of North American stream fishes*. Norman, OK: University of Oklahoma Press, 251–257.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Simons AM. 2004. Phylogenetic relationships in the genus *Erimystax* (Actinopterygii: Cyprinidae) based on the cytochrome b gene. *Copeia* **2004:** 351–356.

Storer DH. 1846. A synopsis of the fishes of North America. *Memoirs of the American Academy of Arts and Sciences* **2:** 253–550.

Strange RM, Burr BM. 1997. Intraspecific phylogeography of North American highland fishes: a test of the Pleistocene vicariance hypothesis. *Evolution* **51:** 885–897.

Sukumaran J, Holder MT. 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* **26:** 1569–1571.

Swofford DL. 2003. *PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods)*. Version 4. Sunderland, MA: Sinauer Associates.

Thomerson JE. 1969. Variation and relationships of the studfishes, *Fundulus catenatus* and *Fundulus stellifer* (Cyprinodontidae, Pisces). *Tulane Studies in Zoology and Botany* **16:** 1–22.

Wiley EO. 1981. *Phylogenetics: the theory and practice of phylogenetic systematics*. Hoboken, NJ: Wiley-Interscience.

Xie WG, Lewis PO, Fan Y, Kuo L, Chen MH. 2011. Improving marginal likelihood estimation for Bayesian phylogenetic model selection. *Systematic Biology* **60:** 150–160.

Yang Z, Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America* **107:** 9264–9269.

Zwickl DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD dissertation, University of Texas at Austin.

SUPPORTING INFORMATION

Additional supporting information may be found online in the supporting information tab for this article:

Figure S1. Resulting gene trees of nuclear DNA introns (*stx5a, ncl1, rpsa, and rps3*) from phylogenetic analysis in a Bayesian framework. Tips represent unique haplotypes. Circles indicate clades supported by ≥ 0.95 Bayesian posterior probabilities.

Table S1. List of species sampled, catalogue number, locality, and GenBank accession number for molecular loci sampled.