Received: 17 December 2019 | Revised: 12 June 2020 | Accepted: 14 June 2020

DOI: 10.1111/zsc.12439

ORIGINAL ARTICLE

Zoologica Scripta | WILEY

Assessing phylogeographic variation in the Rosyside Dace (Teleostei, Leuciscidae), a widespread morphologically variable taxon

Courtney A. Weyand[1,2] | **Kyle R. Piller**[1]

[1]Department of Biological Sciences, Southeastern Louisiana University, Hammond, LA, USA

[2]Department of Biological Sciences, Auburn University, Auburn, AL, USA

Correspondence
Courtney A. Weyand, Department of Biological Sciences, Auburn University, 331 Funchess Hall, Auburn, AL 36849, USA.
Email: courtney.weyand@auburn.edu

Abstract

Clinostomus (Leuciscidae) is a wide-ranging freshwater fish genus that occurs throughout eastern North America and southern portions of Canada with two species currently recognized: *C. elongatus* and *C. funduloides*. A previous taxonomic study of *C. funduloides* recognized two subspecies (*estor* and *funduloides*) and one diagnosed, but undescribed subspecies based on morphological characteristics and geographic distribution. In this study, we used three molecular markers (cytochrome b, S7 intron 1 and growth hormone intron 4) to test the three lineage hypothesis and evaluate genetic variation of *C. funduloides* across the range using Bayesian inference. Our results indicate that *C. funduloides* is not monophyletic, as individuals of *C. elongatus* nest within *C. funduloides* in both the mtDNA and nDNA phylogenetic analyses, although the position of *C. elongatus* varies between data sets. In addition, some of the recovered clades are deeply divergent from one another, further supporting the distinctiveness of many of the populations. Overall, these results suggest that subspecies designations are not warranted and a taxonomic revision is needed as *Clinostomus* is likely more diverse than is currently recognized.

KEYWORDS

phylogeography, Rosyside Dace, systematics

1 | INTRODUCTION

Freshwater fishes in North America are substantially diverse, containing more than 1,100 species (Smith, Badgley, Eiting, & Larson, 2010). The family Leuciscidae is the most species-rich family found in North America, with approximately 59 genera and 322 described species, representing four of the six subfamilies within Leuciscinae (Fricke, Eschmeyer, & Fong, 2019; Schonhuth, Vukic, Sanda, & Mayden, 2018; Tan & Armbruster, 2018). Given the substantial amount of morphological, ecological and species-level diversity within, this family has been of great interest from taxonomic and phylogenetic perspectives (Broughton & Gold, 2000; Bufalino & Mayden, 2010; Cashner, Piller, & Bart, 2011; Martin & Bonett, 2015; Nagle & Simons, 2012; Schonhuth, Shiozawa, Dowling, & Mayden, 2012; Simons & Mayden, 1999).

While many North American leuciscid genera have been heavily studied from both a taxonomic and systematic perspective, the genus *Clinostomus* (Teleostei: Leuciscidae) has been understudied relative to other genera. *Clinostomus* is disjunctly distributed across a large geographic range, occupying most of the eastern North America. It is a member of the open posterior myodome (OPM) clade (Mayden, 1989; Simons, Berendzen, & Mayden, 2003) and is most closely related to two western North American leuciscid genera, *Richardsonious* and *Iotichthys* (Houston, Shiozawa, & Riddle, 2010). Within the genus *Clinostomus*, there are currently two described species: Redside Dace and Rosyside

Dace. The Redside Dace, *Clinostomus elongatus* (Kirtland 1840) (Figure 1), is disjunctly distributed throughout the upper Mississippi River, Great Lakes, Ohio and upper Susquehanna River drainages (COSEWIC, 2017; Novinger & Coon, 2000; Page & Burr, 2011). The other species in the genus, the Rosyside Dace, *Clinostomus funduloides* (Girard, 1856), also has a disjunct distribution, being known from the Atlantic slope from the lower Delaware River drainage south to the Savannah River drainage, and into the continental interior within the Mississippi and Ohio River drainages (Page & Burr, 2011). Two subspecies are currently recognized within *C. funduloides* (Deubler, 1955). The nominal subspecies *C. funduloides funduloides* (Girard, 1856) is distributed along the northern and eastern portions of the range of *C. funduloides* and along the Atlantic slope and upper Ohio River basins with the type locality in the Potomac River basin near Washington, DC. The second recognized subspecies, *Clinostomus funduloides estor* (Jordan & Brayton, 1878), is confined to the lower and middle Cumberland and Tennessee River drainages in Kentucky, Tennessee and Alabama with the type material coming from the Elk River (Tennessee River) and the Stones River (Cumberland River).

In a previous morphological study conducted on *C. funduloides*, Deubler (1955) recognized two subspecies and a diagnosed subspecies that was never formally described in a published work and remains a *nomen nudum* (ICZN, 1999). The morphologically diagnosed, but undescribed, subspecies, commonly known as the 'Smoky Dace' by North American ichthyologists (Jelks et al., 2008; Warren et al., 2000), is restricted to the Little Tennessee River system in Tennessee and North Carolina. Deubler (1955) indicated that the nominal subspecies, *C. f. funduloides,* and the undescribed subspecies ('Smoky Dace') intergrade in the extreme headwaters of the Little Tennessee River system and in the headwaters of the Savannah River drainage in northern Georgia. It has also been noted that the Hiwassee River system, Tennessee

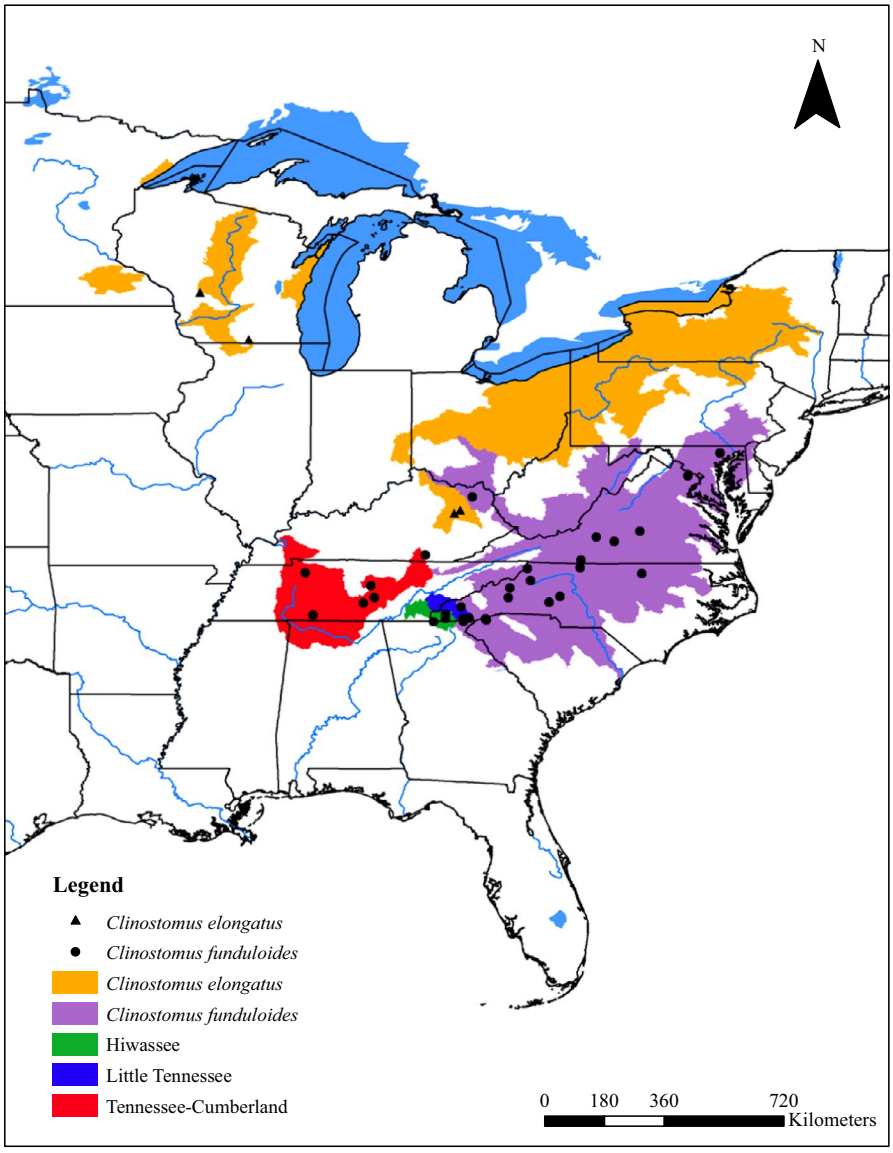

FIGURE 1 Geographic distribution map of species within the genus *Clinostomus*. Colours represent species geographic distributions. The distribution of *Clinostomus elongatus* is coloured orange and the distribution of *Clinostomus funduloides* is coloured purple. The red, blue and green coloured areas represent subspecies/intergrades. Black dots represent collection sites in this study [Colour figure can be viewed at wileyonlinelibrary.com]

and North Carolina, may contain intergrades between *C. f. estor* and the undescribed Smoky Dace (Deubler, 1955; Page & Burr, 2011).

While subspecies are currently recognized as valid taxonomic units (ICZN, 1999), the validity of subspecies is heavily debated among taxonomists since the development of the evolutionary species concept (Wiley, 1981), because subspecies do not represent distinct evolutionary lineages (Burbrink, Lawson, & Slowinski, 2000; Cracraft, 1983; Frost & Hillis, 1990; McKitrick & Zink, 1988). *C. funduloides* is a morphologically variable species; however, the taxonomic validity of its populations has not been tested in a phylogenetic framework. Therefore, the objective of this study was to conduct a comprehensive phylogeographic analysis of *C. funduloides* using mtDNA and nDNA markers, specifically focusing on whether the current taxonomic history is in agreement with the evolutionary history.

2 | MATERIAL AND METHOD

2.1 | Taxon sampling

Specimens of *Clinostomus* were collected from 36 localities across the range using standard seines and electrofishing. Tissue samples were preserved in 95% ethanol, and voucher specimens of these tissues were preserved in 10% formalin. Tissue samples were deposited into the Southeastern Louisiana University Tissue Collection (SLU-TC), and voucher specimens were deposited into the Southeastern Vertebrate Museum (SLU). Other tissue samples were provided by colleagues or natural history museums. Localities are summarized in Table S1. Museum abbreviations follow Sabaj (2016) and hierarchical river terminology follows Jenkins, Lachner, and Schwartz (1972). *Richardsonius balteatus* was included in the analysis as the outgroup, and additional specimens of *C. elongatus* from previously published phylogenies (Houston et al., 2010) were gathered from GenBank.

2.2 | Molecular methods

Whole genomic DNA was extracted from tissue samples using the DNeasy Tissue Kit (Qiagen, Inc.) and run on 0.8% agarose gel to determine the extracted product quality. Three genes, cytochrome b (cytb), S7 intron 1 (S7-1) and a portion of growth hormone intron 4 (GH-IV), were amplified with PCR. Cytochrome b was amplified using primers L12724 (forward) and H15915 (reverse) (Schmidt & Gold, 1993) in 25 μl reactions using the following PCR reaction: initial denaturation at 95°C for 120 s, 30 cycles of denaturation at 95°C for 30 s, annealing at 52°C for 30 s, extension at 72°C for 60 s and final extension at 72°C for 10 min.

S7-1 was amplified using the primers S7RPEX1F (forward) and S7RPEX2R (reverse) (Chow & Hazama, 1998) in 25 μl reactions using the following PCR reaction: initial denaturation at 95°C for 120 s, followed by 1 cycle each at 95°C for 30 s, with annealing stepdown of 1°C increments from 64°C to 56°C for 45 s, and extension 72°C for 60 s, followed by 21 cycles of initial denaturation at 95°C for 30 s, annealing 55°C for 45 s, extension 72°C for 60 s and final extension of 72°C for 5 min.

GH-IV was amplified using nested PCR with four primers identified in Moyer, Remington, & Turner, 2009. Primer set one GHe3.min.3F (forward) and Ghe5.183R (reverse) and primer set two GHe4.min.11F (forward) and Ghe5.173R (reverse). The initial round (primer set one) was performed in 25 μl reaction using the following PCR reaction: initial denaturation at 94°C for 60 s, 30 cycles of denaturation at 94°C for 30 s, annealing 52°C for 30 s, extension 72°C for 90 s and final extension of 72°C for 10 min. The initial PCR product was diluted 1:99 with sterilized H_2O. 1 μl of diluted PCR product was used as the starting template with primer set two in 25 μl reactions using the following PCR reaction: initial denaturation at 94°C for 60 s, 30 cycles of denaturation at 94°C for 30 s, annealing 52°C for 30 s, extension 72°C for 30 s and final extension of 72°C for 10 min.

PCR products for cytb and S7-1 were visualized on a 0.8% agarose gel, and products for GH-IV were visualized on 1.3% and compared to the standard to assess the quality, presence and size of the amplified fragments. Amplified unpurified products for all loci were sent to an external sequencing facility (GENEWIZ, Cambridge, MA) for DNA sequencing. Resulting sequences were edited and aligned manually using the MUSCLE algorithm (Edgar, 2004) as implemented in Geneious 10.1.3 (Kearse et al., 2012) and then submitted to GenBank (Cytb Accession Nos: MT467570-MT467663, S7 Accession Nos: MT490248-MT490301, GH-IV Accession Nos: MT588347-MT588400).

2.3 | Phylogenetic analysis

We included 98 individuals in a phylogenetic analysis of mitochondrial DNA cytochrome b gene, and 54 individuals were used in the nuclear DNA S7 intron 1 and growth hormone intron 4 concatenated phylogenetic analysis. Outgroup taxa included closely related leuciscid (Houston et al., 2010), Redside shiner (*Richardsonius balteatus*).

Using Bayesian inference, phylogenetic relationships within *Clinostomus* for cytb were inferred independently and S7-1 and GH-IV were concatenated. Partition schemes and model selection was performed for the cytb loci and the two concatenated nuclear genes using PartitionFinder v.2.1.1 (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2016) in

order to determine the optimal partitioning scheme and best fit models of nucleotide substitutions. Cytb was partitioned by codon position with models of evolution chosen for each partition using the Bayesian information criterion (BIC). Nuclear markers, S7-1 and GV-IV, were not partitioned due to their non-coding nature; appropriate models of evolution were chosen using the same approach.

Bayesian inference of partitioned loci was conducted in MrBayes 3.2.6 (Huelsenbeck & Ronquist, 2001) on the CIPRES portal science gateway (Miller, Pfeiffer, & Schwartz, 2010). A 50% majority rule consensus tree for cytb and the nuclear concatenated data set was generated by conducting two independent runs for 30,000,000 generations each, sampling every 3,000 generations, with the first 25% of these discarded as burn-in. Bayesian posterior probabilities were calculated using the postburn trees and then visualized in FigTree v.1.4.2 (Rambaut, 2014).

Genetic distance for cytb and concatenated S7-1 and GH-IV was calculated using pairwise uncorrected p distances in MEGA v.7.0.26 (Kumar, Stecher, & Tumura, 2016).

2.4 | Haplotype networks

Haplotype networks based on cytb were generated for the complete data set using PopART v.1.7 (http://popart.otago.ac.nz). To estimate genealogical intraspecific relationships, a TCS network was constructed (Clement, Posada, & Crandall, 2001). The cytb sequences used for the haplotype network analysis were trimmed to the shortest sequence (1,073 bp) to avoid overestimating distinct haplotypes. A complete network with all sequences was first created to determine the number of substitutions between haplotypes and then separated into unique haplotype networks based on their respective localities. Separation of the haplotypes was supported by TCS v1.21 analysis using a 95% cut-off criterion.

3 | RESULTS

3.1 | Sequence alignment and model selection

Sequence data were obtained from 85 individuals of *C. funduloides*, 12 individuals of *C. elongatus* and an outgroup taxon *Richardsonius balteatus* (Table S1). For cytb, sequences were trimmed to make each sequence equal in length. The final length for cytb was 1,073 bp. Individual sequence lengths for both nuclear genes varied in length. The final for S7-1 was 861 bp and 508 bp for GH-IV, with insertions and deletions. The best-fitting partition scheme and model selection for phylogenetic analysis for each gene were as follows: cytb by codon position is F81, GTR + G and K80, S7-1 is HKY, and JC + G for GH-IV.

3.2 | Phylogenetic analysis (mtDNA)

Phylogenetic analysis of the mitochondrial cytb marker recovered two major clades (Figure 2): Clade A, an Atlantic slope, upper Tennessee River and Ohio River clade, and Clade B, representing the Hiwassee, Little Tennessee, Tennessee and Cumberland Rivers, and *C. elongatus*. Both clades were strongly supported with BPP = 100 and interclade (Clade A vs. Clade B) genetic divergence was high at 8.37% uncorrected *p*-distance.

Clade A includes individuals of *C. funduloides* from the Atlantic slope, upper Tennessee and Ohio River basins (Figure 2). The relationships within this clade remain largely unresolved; however, several strongly supported subclades within Clade A were recovered. The first subclade represents specimens from the upper Atlantic slope and upper Ohio River, including samples from the James, Roanoke, Potomac Rivers and Chesapeake Bay on the middle Atlantic slope, and the Green River and Tygarts Creek in the Ohio River basin. The remainder of Clade A is unresolved and includes specimens from the lower Atlantic slope drainage including the Pee Dee, Savannah and Santee Rivers, as well as the French Broad in the Tennessee River basin and the New River in the upper Ohio River basin. Within Clade A, mean uncorrected *p*-distance was 2.17% (0.00%–3.63%).

Clade B includes representatives of *C. funduloides* found outside of the Atlantic slope and Ohio River drainages, as well as multiple individuals of *C. elongatus* sequenced in this study from the middle Ohio River basin of Kentucky and others downloaded from GenBank originally collected from the Upper Mississippi River basin in Wisconsin. Three distinct subclades were recovered. The first subclade represents individuals found within the Hiwassee River from North Carolina and Georgia. This subclade is recovered as monophyletic with a BPP = 100 and is sister to the remaining subclades within Clade B. Average uncorrected sequence divergence between the Hiwassee subclade and remainder of Clade B is high 4.58% (0.00%–5.68%). The second subclade includes individuals found within the Little Tennessee River system. The clade is recovered as an unresolved polytomy, but monophyletic with a BPP = 100 and sister to the remaining portion of the clade. Divergence between the Little Tennessee subclade and the Hiwassee subclade is high 4.30% (0.00%–5.68%). The final subclade represents specimens of *C. funduloides* from the Tennessee and Cumberland Rivers and a monophyletic clade of *C. elongatus*. High genetic divergence is recovered within this subclade 2.87%. Uncorrected *p*-distance values between mtDNA clades and outgroup taxon *R. balteatus* averaged 5.75% (Table 1).

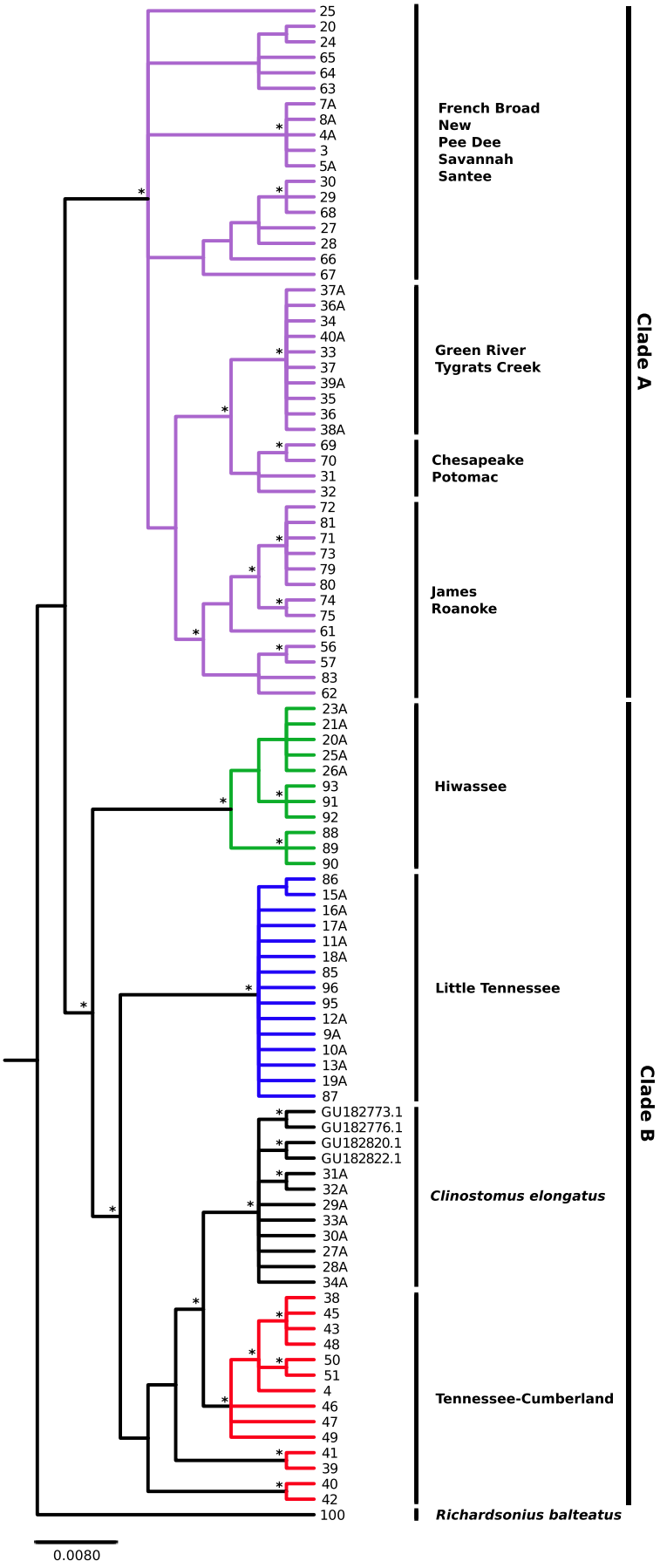

FIGURE 2 Fifty per cent majority rule consensus tree from Bayesian inference of the mitochondrial gene cytb. Nodes labelled with an asterisk (*) indicate posterior probabilities > 95%. Tips are labelled with numbers corresponding to Table S1 [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Uncorrected *p*-distance values calculated in MEGA. Values below the diagonal represent mitochondrial cytb distances

	Atlantic Slope Ohio River	Hiwassee	Little Tennessee	Tennessee–Cumberland	*C.elongatus*	Outgroup
Atlantic Slope Ohio River		0.0276 (2.76%)	0.0234 (2.34%)	0.0258 (2.58%)	0.0270 (2.70%)	0.0332 (3.32%)
Hiwassee	0.0811 (8.11%)		0.0092 (0.92%)	0.0117 (1.17%)	0.0270 (2.70%)	0.0330 (3.30%)
Little Tennessee	0.0837 (8.37%)	0.0434 (4.34%)		0.0075 (0.75%)	0.0228 (2.28%)	0.0288 (2.88%)
Tennessee–Cumberland	0.0842 (8.42%)	0.0416 (4.16%)	0.0350 (3.50%)		0.0253 (2.53%)	0.0313 (3.13%)
C. elongatus	0.0853 (8.53%)	0.0537 (5.37%)	0.0520 (5.20%)	0.0456 (4.56%)		0.0361 (3.61%)
Outgroup	0.0992 (9.92%)	0.0885 (8.85%)	0.0926 (9.26%)	0.0943 (9.43%)	0.0923 (9.23%)	

Note: Values above the diagonal represent nuclear concatenated S7-1 and GH-IV distances.

3.3 | Phylogenetic analysis (nDNA) concatenated

Phylogenetic analysis of the nuclear concatenated S7-1 and GH-IV data set recovered two major clades (Figure 3). Several of the nuclear clades were similar to those recovered in the mitochondrial cytb data set, but there are differences. Clade C includes an unresolved subclade inclusive of individuals from the Atlantic slope, upper Tennessee and Ohio River basins. This subclade is sister to *C. elongatus*, resulting in a paraphyletic *C. funduloides*. Clade D recovered three subclades similar to the cytb data set. A monophyletic Hiwassee River clade BPP = 99, a monophyletic Little Tennessee River clade BPP = 100, and a monophyletic Tennessee–Cumberland River clade BPP = 97, which was recovered as paraphyletic in the mitochondrial cytb dataset. In addition, the Hiwassee River clade is recovered as sister to the Little Tennessee River clade. Overall, the relationships recovered with the nuclear data set are similar to the mitochondrial dataset with the exception of the placement of sister taxon, *C. elongatus* and the interspecific relationships.

Uncorrected *p*-distance values between nDNA clades and outgroup taxon *R. balteatus* averaged 1.93% (0.00%–3.86%) (Table 1).

3.4 | Haplotype networks

Haplotype network using the TCS 95% cut-off criterion for cytb data were separated by 15 mutational steps. A total of 48 unique haplotypes were recovered for 97 individuals, inclusive of *C. elongatus* and separated into nine distinct networks (Figure 4). Networks A-H represent sequences of *C. funduloides* and network I represents sequences of *C. elongatus*. Network A recovered six haplotypes representing individuals from the upper Atlantic slope and upper Ohio River drainages. One haplotype was shared between two localities draining into the Ohio River, Tygarts Creek and Green River. Network B recovered 10 haplotypes representing individuals from the mid-Atlantic slope drainage. Network C recovered 11 haplotypes representing the lower Atlantic slope drainage. Two haplotypes were shared between the Pee Dee and Santee Rivers and the Pee Dee and New Rivers. Network D recovered five haplotypes from the Hiwassee River. Network E recovered four haplotypes from the Little Tennessee River with one haplotype shared among all four sampled localities. Haplotype F recovered two haplotypes from the Tennessee River tributary, Shoal Creek. Haplotype G recovered one haplotype from the Cumberland River tributary, Wells Creek. Haplotype H recovered four haplotypes sequences representing the Tennessee and Cumberland Rivers with two haplotypes shared between Duck River and Caney Fork and McMahan and Dry Creeks. The final haplotype I recovered five haplotypes of *C. elongatus* with one haplotype shared between the Red and Licking Rivers in Kentucky.

4 | DISCUSSION

The subspecies concept has long been a contentious and heavily scrutinized taxonomic category (Wilson & Brown, 1953), but, despite this, has regularly been applied in a variety of different organismal groups. In recent years, however, utilization of trinomials has fallen out of favour in many groups as many researchers view this concept as antiquated and philosophically inaccurate. Many have argued that species, rather than subspecies, represent the true unit of biodiversity (Frost & Hillis, 1990; Zink, 2004). At the time of the original morphological review of *C. funduloides*, the biological species concept (sensu Mayr, 1942) was the predominant concept that guided researcher's taxonomic decisions and lead to the recognition of subspecies, particularly in areas of perceived intergradation, such as the Little Tennessee, Savannah and Hiwassee Rivers, in the case of *C. funduloides*. Incorporation of alternative species concepts (i.e. evolutionary species concept), as well as the inclusion of additional data sources, such as mtDNA and nDNA sequences, has resulted in

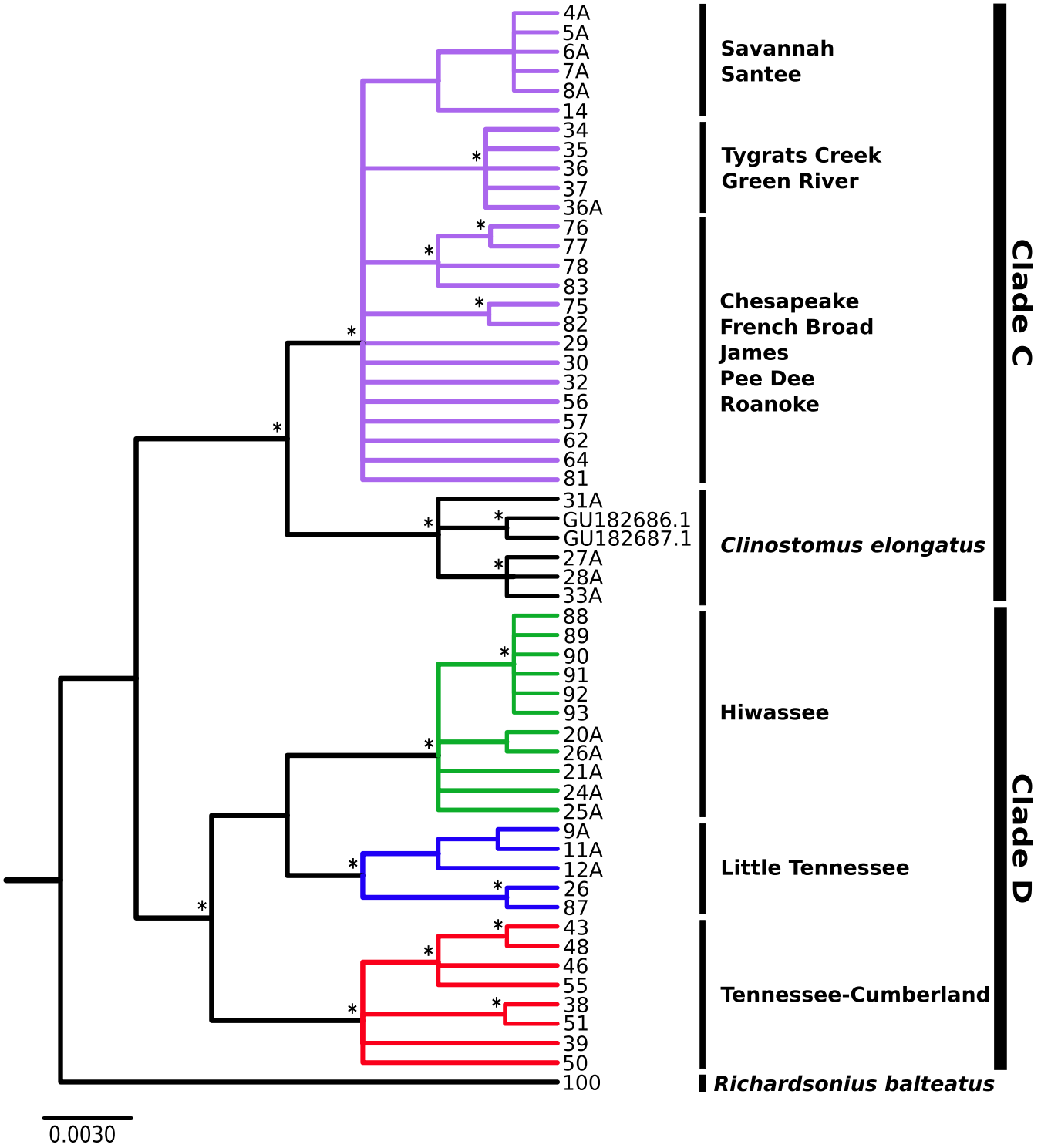

FIGURE 3 Fifty per cent majority rule consensus tree from Bayesian inference of the nuclear concatenated genes S7-1 and GH-IV. Nodes labelled with an asterisk (*) indicate posterior probabilities > 95%. Tips are labelled with numbers corresponding to Table S1 [Colour figure can be viewed at wileyonlinelibrary.com]

different interpretations as to the taxonomic status of subspecies in many different groups (Hayes & Piller, 2018; Kraczkowski & Chernoff, 2014; Martin et al., 2013; Piller & Bart, 2017).

In the case of *C. funduloides*, the taxonomic status of many of the populations has been in flux. This is due, in part, to its disjunct distribution across eastern North America and the known morphological variation within this species (Deubler, 1955). This includes the recognition of two subspecies (*C. f. funduloides* and *C. f. estor*) (Deubler, 1955), as well as the recognition of additional undescribed forms (i.e. Smoky Dace) (Jelks et al., 2008; Warren et al., 2000). Prior to

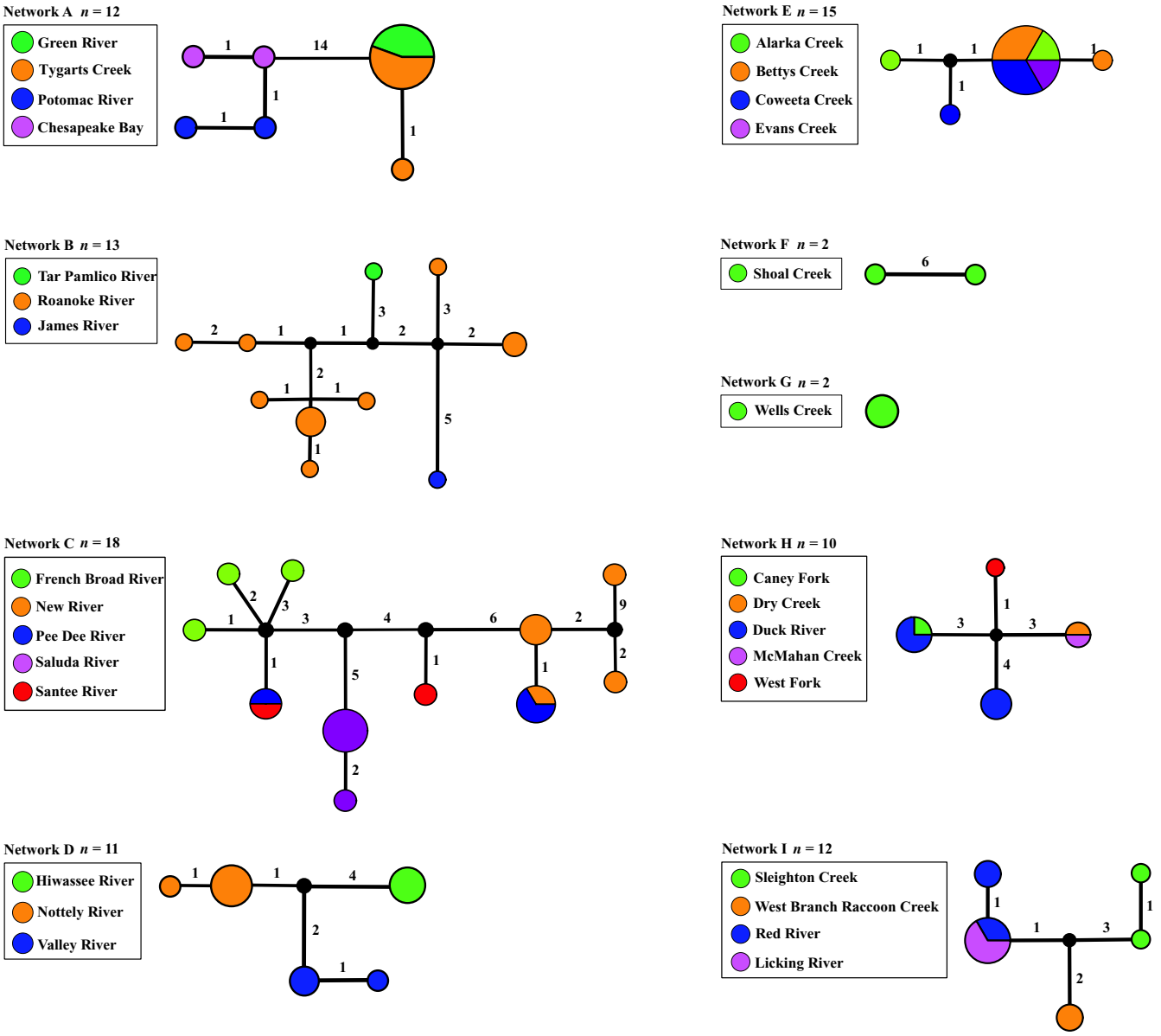

FIGURE 4 TCS haplotype networks of the mitochondrial gene cytb. Circle size is proportional to the number of samples within a given haplotype and number between haplotypes represent mutational steps between. Colours represent specific localities. Haplotype networks A-H represent *Clinostomus funduloides* and haplotype network I represents *Clinostomus elongatus* [Colour figure can be viewed at wileyonlinelibrary.com]

this study, no study has quantified molecular variation across the range of *C. funduloides*.

4.1 | Mitochondrial DNA

Phylogenetic analysis of cytb (mtDNA) recovered two major clades of *Clinostomus*. Clade A included individuals found within the Atlantic slope, Ohio River and upper Tennessee River basins, and there was strong support for this clade. Interclade relationships, however, were weakly supported and relatively unresolved. The populations within Clade A mostly agree with Deubler (1955) with the exception of the Savannah River population. Deubler (1955) identified this population an intergrade between *C. f. funduloides* and the Little Tennessee River population of *C. funduloides*. As a result, these populations were not assigned to any particular subspecies. In this present study, the Savannah River samples grouped with the Atlantic slope, upper Tennessee and Ohio River clade (Clade A) rather than the Little Tennessee samples (Clade B) as proposed by Deubler (1955). Our data demonstrate that both populations possess unique haplotypes and are 8.11% divergent from one another. A close relationship between the Atlantic slope and Ohio River has been documented for many other aquatic species: *Etheostoma caeruleum* (Esmond & Stauffer, 1983), *Notropis rubellus* (Berendzen, Olson, & Barron, 2009) and the genus *Nocomis* (Nagle & Simons, 2012).

The second major clade recovered with the cytb data set, Clade B, contained individuals of *C. funduloides* from outside of the Atlantic slope and Ohio River ranges as well as its sister taxon, *C. elongatus*. A subclade representing individuals from the Hiwassee River is recovered as monophyletic. Deubler (1955) suggested that the Hiwassee River population represented an intermediate form between *C. f. estor* and a diagnosed, but undescribed subspecies from the Little Tennessee River. Our results, however, suggest that the Hiwassee River population is its own independent lineage. This population is monophyletic and genetically divergent, more than 4% (uncorrected *p*-distance) on average, from the other subclades within clade B. The uniqueness of this lineage is further supported by the existence of its own unconnected haplotype network and the possession of unique haplotypes.

A second unique, monophyletic subclade, the Little Tennessee River subclade within Clade B, was recovered. Resolution within it was poor, suggesting little genetic diversity within the group. The haplotype analysis recovered four haplotypes with a single common haplotype and three singleton haplotypes. Furthermore, the Little Tennessee River subclade is genetically divergent (>3.5% on average) from other subclades within Clade B. The genetic distinctiveness of the subclade also is not surprising, as this river system possesses a suite of unique and endemic freshwater fishes. This subclade corresponds to the undescribed subspecies identified by Deubler (1955).

Finally, the most striking result is the close relationship of the Tennessee–Cumberland River subclade of *C. funduloides* and *C. elongatus*. This relationship has not previously been recovered or suggested, as most other phylogenetic studies failed to include the diversity of *Clinostomus* samples as in this present study (Bufalino & Mayden, 2010; Houston et al., 2010; Schonhuth et al., 2012). Morphologically, there has been little question in regard to the taxonomic identity of the two species of *Clinostomus*. Both species are easily identifiable, as *C. elongatus* differs from *C. funduloides* in mouth size, snout shape, lateral line scale counts and many other characters (Page & Burr, 2011). The recovered topology renders *C. funduloides* as paraphyletic. Several of our samples of *C. elongatus* (27A-34A) were collected near populations of *C. funduloides* from Kentucky. However, inclusion of additional samples from GenBank (GU182773.1, 182776.1, 182820.1 and 182822.1), from Wisconsin, resulted in the recovery of the same topology and paraphyly for *C. funduloides*.

The TCS haplotype networks for cytb support similar groupings as the phylogenetic analysis. Twenty-seven haplotypes were present within the Atlantic slope–Ohio River in three disconnected networks. Within the network, three haplotypes were shared among, suggesting that these river systems are experiencing gene flow across the geographic extent of the clade. Five unique haplotypes were found within the Hiwassee River system. Four unique haplotypes were found within the Little Tennessee, and six haplotypes were present within the Tennessee–Cumberland River clade in three disconnected networks.

4.2 | Nuclear DNA

The nDNA concatenated data set of S7-1 and GH-IV also revealed two major clades (Clade C and Clade D, Figure 3). Clade C recovered a strongly supported clade of Atlantic slope, Ohio and Upper Tennessee River populations that correspond to *C. f. funduloides*. Like the mtDNA, the Savannah River and upper Tennessee River populations were recovered within a larger Atlantic slope, Ohio River and upper Tennessee River clade. In addition, samples of *C. elongatus* were recovered as sister to a group of individuals of *C. funduloides* from the Atlantic slope, Ohio and Upper Tennessee Rivers, and this differs from results of the mtDNA analysis, which grouped *C. elongatus* with the Tennessee–Cumberland River clade.

Clade D recovered relationships similar to the mtDNA data set; however, the Tennessee and Cumberland Rivers clade is recovered as monophyletic, with strong support. This suggests that *C. f. estor* represents a distinctive lineage that may warrant species-level taxonomic status. Furthermore, the Hiwassee and Little Tennessee River clades are each recovered as monophyletic. Other studies have identified unique aquatic lineages of other species from the Hiwassee River (Cooper, 2006; Hobbs, 1981), Little Tennessee River (Blanton & Jenkins, 2008; Dinkins & Shute, 1996) or both (Piller, Bart, & Hurley, 2008). The recovery of distinctive Hiwassee and Little Tennessee River clades also indicates that these lineages are in need of additional morphological examination, as they likely are specifically distinct.

4.3 | Mitochondrial and nuclear DNA discordance

Discordance among topologies generated from mtDNA and nDNA markers is not an uncommon phenomenon (Bossu & Near, 2009; Sota & Vogler, 2001; Toews & Brelsford, 2012). These markers typically have different rates of molecular evolution and effective population sizes, both of which can impact the recovered topologies (Ballard & Whitlock, 2004). This is particularly true for recently diverged groups, such as *Clinostomus*, in which the two currently recognized species, *C. elongatus* and *C. funduloides*, are estimated to have diverged approximately 2.6 mya (Houston et al., 2010). Despite the discordance, the same general lineages were recovered with both data sets.

5 | CONCLUSION

Using multiple loci, the phylogeographic relationships among the populations of *Clinostomus* were inferred. Results based on mtDNA and nDNA sequences indicate that there is a greater amount of species-level diversity within the genus than is currently recognized. The taxonomic history of a group of organisms should be reflective of its evolutionary history (Wiley, 1981), and the results from this study suggest that a taxonomic revision of *Clinostomus* is needed. Based on meristics, proportional measurements and pigmentation differences, Deubler (1955) noted that *C. funduloides* consists of three geographic subspecies and intergrades zones. The molecular data generated in this study generally support the taxonomic groups recognized by Deubler (1955); however, it also supports the existence of additional distinctive lineages. The identification of distinctive molecular lineages, in the Tennessee–Cumberland, Atlantic slope–upper Tennessee–Ohio Rivers, Hiwassee River and Little Tennessee River and the general concordance of independent genomes provides additional support for a taxonomic revision. Furthermore, both mtDNA and nDNA data sets recover *C. funduloides* as paraphyletic and in both cases, *C. elongatus* is more closely related to a subclade of *C. funduloides* rather than the *C. funduloides* subclades are to each other, although the position of *C. elongatus* differs between mtDNA and nDNA data sets. We refrain from proposing new names and resolving the taxonomy of *Clinostomus* as future work should utilize the phylogenetic hypotheses generated in the study as an historical template to re-examine the morphological variation within *Clinostomus* and diagnose and re-diagnose the taxa within this interesting and diverse clade of leuciscid fishes.

ACKNOWLEDGEMENTS

We would like to thank multiple curators, collection managers and colleagues who provided tissue samples, David Werneke at Auburn University Museum of Natural History, Mollie F. Cashner and Rebecca Blanton Johansen at Austin Peay State University, David Eisenhour at Morehead State College, Gabriela Houge at North Carolina State Museum, Brian Sidlauskas at Oregon State Museum, Steve Powers at Roanoke College, Ben Keck at the University of Tennessee, Marc Kibbey at Ohio State University, Dave Neely at Tennessee Aquarium Conservation Institute, Matt Thomas and Stephanie Brandt at Kentucky Department of Fish and Wildlife, and John Lyons at the University of Wisconsin-Madison for providing specimens. We also thank Dr. Katelyn Lawson from the Alabama Natural Heritage Program/Auburn University Museum of Natural History for providing assistance with GIS/mapping. Additionally, we thank D. Cooper Campbell and C. Elyse Parker for providing assistance in the field. This study was supported, in part, by a Schlieder Professorship to KRP.

ORCID

Courtney A. Weyand https://orcid.org/0000-0001-7162-2462
Kyle R. Piller https://orcid.org/0000-0003-1289-9351

REFERENCES

Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology*, *13*, 729–744. https://doi.org/10.1046/j.1365-294X.2003.02063.x

Berendzen, P. B., Olson, W. M., & Barron, S. M. (2009). The utility of molecular hypotheses for uncovering morphological diversity in the *Notropis rubellus* species complex (Cypriniformes: Cyprinidae). *Copeia*, *2009*, 661–673. https://doi.org/10.1634/CI-09-013

Blanton, R. E., & Jenkins, R. E. (2008). Three new darter species of the *Etheostoma percnurum* species complex (Percidae, subgenus *Catonotus*) from the Tennessee and Cumberland river drainages. *Zootaxa*, *1963*, 1–24. https://doi.org/10.11646/zootaxa.1963.1.1

Bossu, C. M., & Near, T. J. (2009). Gene trees reveal repeated instances of mitochondrial DNA Introgression in orangethroat darters (Percidae: Etheostoma). *Systematic Biology*, *58*, 114–125. https://doi.org/10.1093/sysbio/syp014

Broughton, R. E., & Gold, J. R. (2000). Phylogenetic relationships in the North American cyprinid genus *Cyprinella* (Actinopterygii: Cyprinidae) based on sequences of the mitochondrial ND2 and ND4 genes. *Copeia*, *2000*, 1–10. https://doi.org/10.1643/0045-8511(2000)2000[0001:PRITNA]2.0.CO;2

Bufalino, A. P., & Mayden, R. L. (2010). Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. *Molecular Phylogenetics and Evolution*, *55*, 274–283. https://doi.org/10.1016/j.ympev.2009.12.017

Burbrink, F. T., Lawson, R., & Slowinski, J. B. (2000). Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): A critique of the subspecies concept. *Evolution*, *54*, 2107–2118. https://doi.org/10.1554/0014-3820(2000)054[2107:MDPOTP]2.0.CO;2

Cashner, M. F., Piller, K. R., & Bart, H. L. (2011). Phylogenetic relationships of the North American cyprinid subgenus *Hydrophlox*. *Molecular Phylogenetics and Evolution*, *59*, 725–735. https://doi.org/10.1016/j.ympev.2011.03.019

Chow, S., & Hazama, K. (1998). Universal PCR primers for S7 ribosomal protein gene introns in fish. *Molecular Ecology*, *7*, 1255–1256. https://doi.org/10.1046/j.1365-294x.1998.00406.x

Clement, M., Posada, D., & Crandall, K. A. (2001). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Cooper, J. E. (2006). A new crayfish of the genus *Cambarus* Erichson, 1846, subgenus *Puncticambarus* Hobbs, 1969 (Decapoda: Cambaridae), from the Hiwassee River basin of North Carolina. *Proceedings of the Biological Society of Washington*, *119*, 81–90. https://doi.org/10.2988/0006-324X(2006)119[81:ANCOTG]2.0.CO;2

COSEWIC (2017). *COSEWIC assessment and status report on the Redside Dace* Clinostomus elongatus *in Canada*. Ottawa, ON: Committee on the Status of Endangered Wildlife in Canada, Xii+63 pp. https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/cosewic-assessments-status-reports/redside-dace-2017.html#toc5

Cracraft, J. (1983). Species concepts and speciation analysis. In R. F. Johnston (Ed.), *Current ornithology*, vol 1 (pp. 159–187). Boston, MA: Springer. https://doi.org/10.1007/978-1-4615-6781-3_6

Deubler, E. E. (1955). *A taxonomic study of the cyprinid fish* Clinostomus vandoisulus *(Valenciennes) in the eastern United States*. Unpublished Doctoral Dissertation. Ithaca, NY: Cornell University.

Dinkins, G. R., & Shute, P. W. (1996). Life histories of *Noturus baileyi* and *N. flavipinnis* (Pisces: Ictaluridae), two rare madtom catfishes in Citico Creek, Monroe County, Tennessee. *Bulletin Alabama Museum of Natural History*, *18*, 43–69.

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*, 1792–1797. https://doi.org/10.1093/nar/gkh340

Esmond, E. F., & Stauffer, J. R. (1983). Taxometric comparison of the Atlantic Slope and Ohio River Populations of *Etheostoma caeruleum* Storer. *The American Midland Naturalist*, *109*, 390–397. https://doi.org/10.2307/2425420

Fricke, R., Eschmeyer, W. N., & Fong, J. D. (2019). Eschmeyer's Catalog of Fishes: Species By Family/Subfamily. http://researcharchive.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp.

Frost, D. R., & Hillis, D. M. (1990). Species in concept and practice: Herpetological Applications. *Herpetologica*, *46*, 86–104.

Girard, C. F. (1856). Researches upon the cyprinoid fishes inhabiting the fresh waters of the United States of America, west of the Mississippi Valley, from specimens in the museum of the Smithsonian Institution. *Proceedings of the Academy of Natural Sciences of Philadelphia*, *8*, 165–213.

Hayes, M. M., & Piller, K. R. (2018). Patterns of diversification in a North American endemic fish, the Blackbanded darter (Perciformes, Percidae). *Zoologica Scripta*, *47*, 477–485. https://doi.org/10.1111/zsc.12288

Hobbs, H. H. Jr (1981). The Crayfishes of Georgia. *Smithsonian Contributions to Zoology*, (318): 1–549. https://doi.org/10.5479/si.00810282.318

Houston, D. D., Shiozawa, D. K., & Riddle, B. R. (2010). Phylogenetic relationships of the western North American cyprinid genus *Richardsonius*, with an overview of phylogeographic structure. *Molecular Phylogenetics and Evolution*, *55*, 259–273. https://doi.org/10.1016/j.ympev.2009.10.017

Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, *17*, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754

International Commission of Zoological Nomenclature [ICZN] (1999). *International code of zoological nomenclature* (4th edn). London: International Trust for Zoological Nomenclature.

Jelks, H. L., Walsh, S. J., Burkhead, N. M., Contreras-Balderas, S., Diaz-Pardo, E., Hendrickson, D. A., … Warren, M. L., (2008). Conservation Status of imperiled North American freshwater and diadromous fishes. *Fisheries*, *33*, 372–407. https://doi.org/10.1577/1548-8446-33.8.372

Jenkins, R. E., Lachner, E. A., & Schwartz, F. J. (1972). Fishes of the Central Appalachian drainages: Their distributions and dispersal. In P. C. Holt (Ed.), *The Distributional History of the Biota of the Southern Appalachians* (pp. 43–117). Blacksburg, VA: Research Division Monograph, Virginia Polytechnic Institute and State University.

Jordan, D. S., & Brayton, A. W. (1878). Contributions to North American ichthyology. Based primarily on the collections of the United States National Museum, vol. III, pt. A: On the distribution of the fishes of the Alleghany region of South Carolina, Georgia, and Tennessee, with descriptions of new or little known species. *Bulletin of the United States National Museum*, *12*, 1–237.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kraczkowski, M. L., & Chernoff, B. (2014). Molecular phylogenetics of the Eastern and Western Blacknose Dace, *Rhinichthys atratulus* and *R. obtusus* (Teleostei: Cyprinidae). *Copeia*, *2014*, 325–338. https://doi.org/10.1643/CG-14-002

Kumar, S., Stecher, G., & Tumura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2016). PartitionFinder 2: New methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, *34*, 772–773. https://doi.org/10.1093/molbev/msw260

Martin, B. T., Bernstein, N. P., Birkhead, R. D., Koukl, J. F., Mussmann, S. M., & Placyk, J. S. Jr (2013). Sequence-based molecular phylogenetics and phylogeography of the American box turtles (*Terrapene* spp.) with support from DNA barcoding. *Molecular Phylogenetics and Evolution*, *68*, 119–134. https://doi.org/10.1016/j.ympev.2013.03.006

Martin, S. D., & Bonett, R. M. (2015). Biogeography and divergent patterns of body size disparification in North American minnows. *Molecular Phylogenetics and Evolution*, *93*, 17–28. https://doi.org/10.1016/j.ympev.2015.07.006

Mayden, R. L. (1989). Phylogenetic studies of North American minnows, with emphasis on the genus *Cyprinella* (Teleostei: Cypriniformes). *Miscellaneous Publications of the Museum of Natural History, University of Kansas*, *80*, 1–189. https://doi.org/10.5962/bhl.title.5480

Mayr, E. (1942). *Systematics and the Origin of Species*. New York, NY: Columbia University Press.

McKitrick, M. C., & Zink, R. M. (1988). Species Concepts in Ornithology. *The Condor*, *90*, 1–14. https://doi.org/10.2307/1368426

Miller, M. A., Pfeiffer, W., & Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE)* (pp. 1–8). New Orleans, LA.

Moyer, G. R., Remington, R. L., & Turner, T. F. (2009). Incongruent gene trees, complex evolutionary processes, and the phylogeny of a group of North American minnows (*Hybognathus* Agassiz 1855). *Molecular Phylogenetics and Evolution*, *50*, 514–525. https://doi.org/10.1016/j.ympev.2008.11.002

Nagle, B. C., & Simons, A. M. (2012). Rapid diversification in the North American minnow genus *Nocomis*. *Molecular Phylogenetics and Evolution*, *63*, 639–649. https://doi.org/10.1016/j.ympev.2012.02.013

Novinger, D. C., & Coon, T. G. (2000). Behavior and physiology of the Redside Dace, *Clinostomus elongatus*, a threatened species in Michigan. *Environmental Biology of Fishes*, *57*(3), 315–326. https://doi.org/10.1023/A:1007526414384

Page, L. M., & Burr, B. M. (2011). *Peterson field guide to freshwater fishes of North America North of Mexico*. Boston, NY: Houghton Mifflin Harcourt.

Piller, K. R., & Bart, H. L. (2017). Rediagnosis of the Tuckasegee Darter, *Etheostoma gutselli* (Hildebrand), a Blue Ridge Endemic. *Copeia*, *105*, 569–574. https://doi.org/10.1643/CI-17-578

Piller, K. R., Bart, H. L., & Hurley, D. L. (2008). Phylogeography of the Greenside Darter complex, *Etheostoma blennioides* (Teleostomi: Percidae): A wide-ranging polytypic taxon. *Molecular Phylogenetics and Evolution*, *46*, 974–985. https://doi.org/10.1016/j.ympev.2007.11.023

Rambaut, A. (2014). *FigTree v1.4.2: Tree figure drawing tool*. Edinburgh: Institute of Evolutionary Biology, University of Edinburgh.

Sabaj, M. H. (2016). *Standard symbolic codes for institutional resource collections in herpetology and ichthyology: An online reference*. Washington, DC: American Society of Ichthyologists and Herpetologists. http://www.asih.org/

Schmidt, T. R., & Gold, J. R. (1993). Complete sequence of the mitochondrial cytochrome b gene in the Cherryfin Shiner, *Lythrurus roseipinnis* (Teleostei: Cyprinidae). *Copeia*, *1993*, 880–883. https://doi.org/10.2307/1447258

Schonhuth, S., Shiozawa, D. K., Dowling, T. E., & Mayden, R. L. (2012). Molecular systematics of western North American cyprinids (Cypriniformes: Cyprinidae). *Zootaxa*, *3586*, 281–303. https://doi.org/10.11646/zootaxa.3586.1.27

Schonhuth, S., Vukic, J., Sanda, R., & Mayden, R. L. (2018). Phylogenetic relationships and classification of the Holarctic family Leuciscidae (Cypriniformes: Cyprinoidei). *Molecular Phylogenetics and Evolution*, *127*, 781–799. https://doi.org/10.1016/j.ympev.2018.06.026

Simons, A. M., Berendzen, P. B., & Mayden, R. L. (2003). Molecular systematics of North American phoxinin genera (Actinopterygii: Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. *Zoological Journal of the Linnean Society*, *139*, 63–80. https://doi.org/10.1046/j.1096-3642.2003.00076.x

Simons, A. M., & Mayden, R. L. (1999). Phylogenetic relationships of North American cyprinids and assessment of homology of the Open Posterior Myodome. *Copeia*, *1999*, 13–21. https://doi.org/10.2307/1447380

Smith, G. R., Badgley, C., Eiting, T. P., & Larson, P. S. (2010). Species diversity gradients in relation to geological history in North American freshwater fishes. *Evolutionary Ecology Research*, *12*, 693–726.

Sota, T., & Vogler, A. P. (2001). Incongruence of mitochondrial and nuclear gene tress in the Carabid Beetles *Ohomopterus*. *Systematic Biology*, *50*, 35–59.

Tan, M., & Armbruster, J. W. (2018). Phylogenetic classification of extant genera of fishes of the order Cypriniformes (Teleostei: Ostariophysi). *Zootaxa*, *4476*, 6–39. https://doi.org/10.11646/zootaxa.4476.1.4

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*, 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Warren, M. L., Burr, B. M., Walsh, S. J., Bart, H. L., Cashner, R. C., Etnier, D. A., … Starnes, W. C. (2000). Diversity, distribution, and conservation status of the native freshwater fishes of the southern United States. *Fisheries*, *25*, 7–31. https://doi.org/10.1577/1548-8446(2000)025<0007:DDACSO>2.0.CO;2

Wiley, E. O. (1981). *Phylogenetics: The theory and practice of phylogenetic systematics*. New York, NY: John Wiley and Sons.

Wilson, E. O., & Brown, W. L. (1953). The subspecies concept and its taxonomic application. *Systematic Biology*, *2*, 97–111. https://doi.org/10.2307/2411818

Zink, R. M. (2004). The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1539), 561–564. https://doi.org/10.1098/rspb.2003.2617

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Weyand CA, Piller KR. Assessing phylogeographic variation in the Rosyside Dace (Teleostei, Leuciscidae), a widespread morphologically variable taxon. *Zool Scr*. 2020;49:563–574. https://doi.org/10.1111/zsc.12439