








## RESEARCH ARTICLE

# Genetic and morphological characterization of the freshwater mussel clubshell species complex (*Pleurobema clava* and *Pleurobema oviforme*) to inform conservation planning

Cheryl L. Morrison<sup>1</sup>  | Nathan A. Johnson<sup>2</sup>  | Jess W. Jones<sup>3</sup>  |  
Michael S. Eackles<sup>1</sup>  | Aaron W. Aunins<sup>1</sup>  | Daniel B. Fitzgerald<sup>1</sup>  |  
Eric M. Hallerman<sup>4</sup>  | Tim L. King<sup>1,†</sup>

<sup>1</sup>U.S. Geological Survey, Eastern Ecological Science Center, Leetown Research Laboratory, Kearneysville, West Virginia, USA

<sup>2</sup>U.S. Geological Survey, Wetland and Aquatic Research Center, Gainesville, Florida, USA

<sup>3</sup>U.S. Fish and Wildlife Service, Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

<sup>4</sup>Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, USA

## Correspondence

Cheryl L. Morrison, U.S. Geological Survey, Eastern Ecological Science Center, Leetown Research Laboratory, Kearneysville, WV 25430, USA.  
Email: cmorrison@usgs.gov

## Funding information

U.S. Fish and Wildlife Service; U.S. Geological Survey

## Abstract

The shell morphologies of the freshwater mussel species *Pleurobema clava* (federally endangered) and *Pleurobema oviforme* (species of concern) are similar, causing considerable taxonomic confusion between the two species over the last 100 years. While *P. clava* was historically widespread throughout the Ohio River basin and tributaries to the lower Laurentian Great Lakes, *P. oviforme* was confined to the Tennessee and the upper Cumberland River basins. We used two mitochondrial DNA (mtDNA) genes, 13 novel nuclear DNA microsatellite markers, and shell morphometrics to help resolve this taxonomic confusion. Evidence for a single species was apparent in phylogenetic analyses of each mtDNA gene, revealing monophyletic relationships with minimal differentiation and shared haplotypes. Analyses of microsatellites showed significant genetic structuring, with four main genetic clusters detected, respectively, in the upper Ohio River basin, the lower Ohio River and Great Lakes, and upper Tennessee River basin, and a fourth genetic cluster, which included geographically intermediate populations in the Ohio and Tennessee river basins. While principal components analysis (PCA) of morphometric variables (i.e., length, height, width, and weight) showed significant differences in shell shape, only 3% of the variance in shell shape was explained by nominal species. Using Linear Discriminant and Random Forest (RF) analyses, correct classification rates for the two species' shell forms were 65.5% and 83.2%, respectively. Random Forest classification rates for some populations were higher; for example, for North Fork Holston (HOLS), it was >90%. While nuclear DNA and shell morphology indicate that the HOLS population is strongly differentiated, perhaps indicative of cryptic biodiversity, we consider the presence of a single widespread species the most likely biological scenario for many of the

<sup>†</sup>Deceased.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd. This article has been contributed to by US Government employees and their work is in the public domain in the USA.

investigated populations based on our mtDNA dataset. However, additional sampling of *P. oviforme* populations at nuclear loci is needed to corroborate this finding.

#### KEYWORDS

COI, cryptic biodiversity, endangered species, microsatellite DNA, mitochondrial DNA, morphometrics, NDI

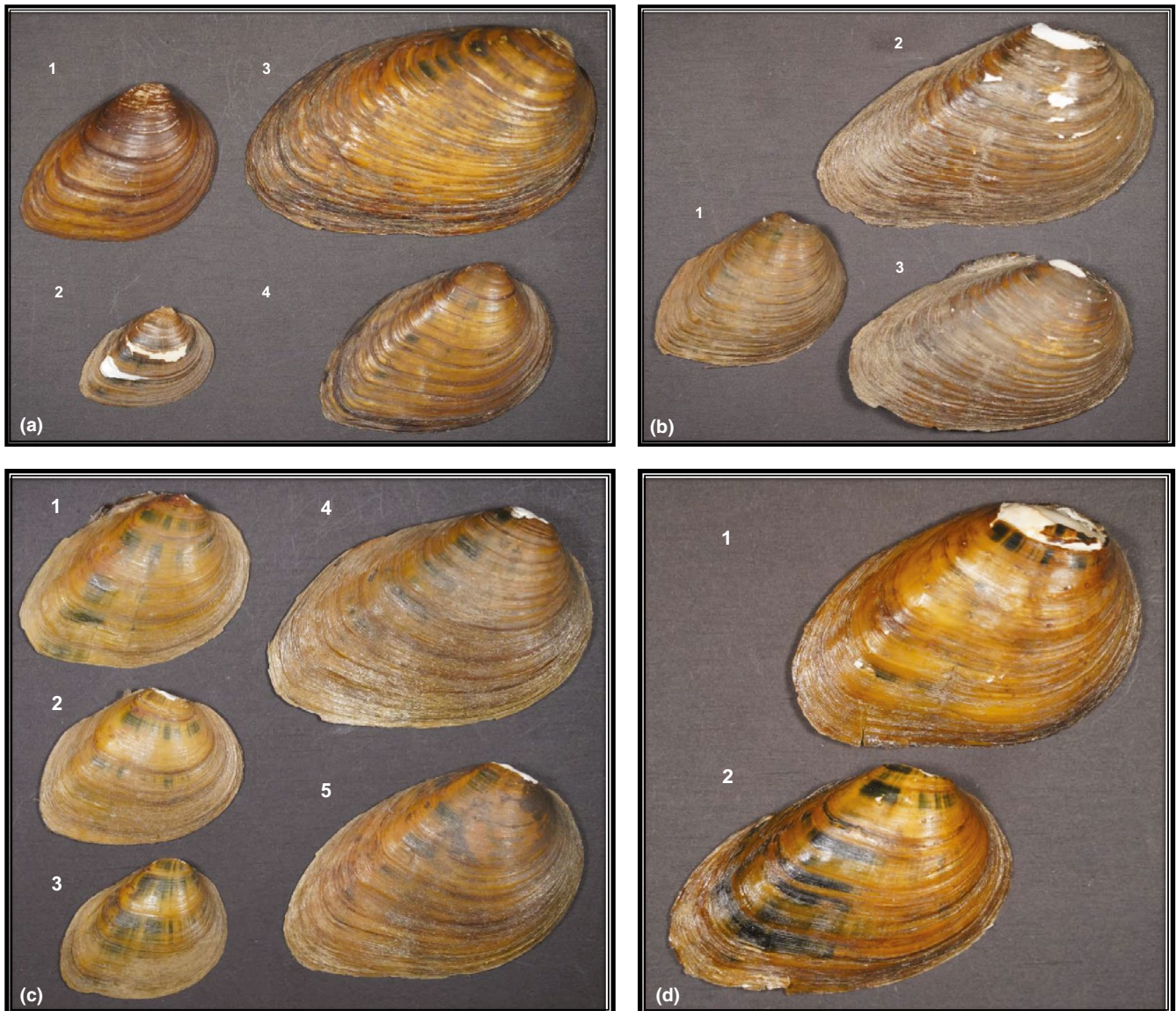
## 1 | INTRODUCTION

A robust scientific basis for managing biodiversity originates with accurate delineation of species and population groups with independent evolutionary trajectories (Bernatchez & Wilson, 1998; De Queiroz, 2007; Pante et al., 2015; Vignieri et al., 2006). However, morphology-based taxonomy does not always reflect evolutionary trajectories (Avice, 1994). While the most diverse in North America, native freshwater mussels (Bivalvia: Unionida) are among the most imperiled groups of animals (Haag, 2012; Lydeard et al., 2004; Williams et al., 1993). Many species inhabit small geographic ranges within stream networks and have limited (and often unknown) dispersal abilities due to reliance on specific host fishes for metamorphosis from larval (glochidia) to free-living life stages (juvenile, sub-adult, adult), making populations easily fragmented and particularly susceptible to habitat modification, pollution, and overharvest (Haag & Williams, 2014; Neves, 1993; Williams et al., 1993). Despite the recognition that many native mussels face high extinction and imperilment rates (Haag & Williams, 2014), knowledge of basic biology, ecology, and taxonomy for some unionids remains lacking, which limits the ability of natural resource managers to take immediate conservation action. Identifying discrete morphological characters to diagnose species or determine evolutionary lineages can be challenging due to similarity in overall shape and appearance. Further, conchological features are heritable, but may also be influenced by ecological factors (Agrell, 1948; Kodolova & Logvinenko, 1973; Ortmann, 1920; Watters, 1994) and can exhibit extensive morphological variation within and among river drainages (Graf, 1997, 1998; Hyde et al., 2020; Inoue et al., 2013, 2015; Johnson et al., 2018; Ortmann, 1920). Such phenotypic plasticity often hinders our ability to delimit species and design appropriate management plans (Burlakova et al., 2012; Inoue et al., 2013, 2014, 2018; Johnson et al., 2018).

Many species belonging to the North American Tribe Pleurobemini (Bivalvia: Unionida) exhibit substantial variation in shell morphology, which has led to taxonomic confusion and reliance on geography to delineate species boundaries. In fact, species belonging to the genus *Pleurobema* have been described as “the most perplexing group of unionids” in some basins (Williams et al., 2008), with 23 species currently recognized (Williams et al., 2017). Most *Pleurobema* species are threatened or endangered at either the state or national level, so difficulty delineating species using morphological characters has strong potential to affect conservation and management efforts (Shea et al., 2011).

The Clubshell, *Pleurobema clava* (Lamarck, 1819), is a freshwater bivalve that inhabits clean, coarse sand, and gravel runs in small- to moderate-sized streams in the eastern United States. *Pleurobema clava* was historically widespread and abundant in the Ohio River basin and tributaries of western Lake Erie (Dean, 1890; Watters, 1988) yet has experienced a 95% range reduction after being found in only 13 of 100 historically known localities (U.S. Fish & Wildlife Service, 1994). *Pleurobema clava* was listed as an endangered species in 1993 (U.S. Fish and Wildlife Service, 1993) under the Endangered Species Act (ESA 1973, as amended). Little is known about the ecology of *P. clava*, or about the connectivity between the remaining, widely scattered populations, and no genetic information exists on population structure, levels of gene flow, or relatedness within or among populations of *P. clava*.

The relationship and taxonomic status between *P. clava* and another currently recognized species, the Tennessee Clubshell, *P. oviforme* (Conrad, 1834), remains uncertain. Past taxonomic confusion between the two species likely centers on the morphological similarity between the majority of *P. oviforme* populations, and the congener *P. clava*. *Pleurobema clava* is elongate or “club-like” (*clava* = “club” in Latin) in appearance (Figure 1a–d), with umbos set-back anteriorly, inflated and extending 2–4 mm beyond the hinge line, whereas *P. oviforme* is oval or “egg-like” (*oviforme* = “egg-form” in Latin) in appearance, with umbos more centrally positioned and much less inflated, typically either even with the hinge line or barely extending beyond it (Figure 2a–f; Parmalee & Bogan, 1998; Watters et al., 2009; Williams et al., 2008). Specifically, shells of *P. oviforme* from the middle (e.g., Paint Rock River) and lower (e.g., Duck River) sections of the Tennessee River basin have inflated umbos that extend well beyond the hinge line, greater width, and tend to be more elongate and have an angular posterior ridge, similar to *P. clava* from the Great Lakes region and Ohio River system. In the Cumberland River basin, however, *P. oviforme* is thought to replace *P. clava* in headwater streams because shells tend to be more compressed and less elongate and have an angular posterior ridge, with umbos that are located anteriorly and extending beyond the hinge line (Figure 2d and e) (Haag & Cicerello, 2016; Parmalee & Bogan, 1998; Watters et al., 2009; Williams et al., 2008). However, when shells of *P. clava* are compared visually to shells of *P. oviforme* from the upper Tennessee River basin, more pronounced differences in shape are evident. The shells of *P. oviforme* from the upper Tennessee River basin are more flattened or compressed (Figure 2f), and the posterior ridge is more rounded and not sharp or angular. These morphological differences allow easy identification of specimens obtained



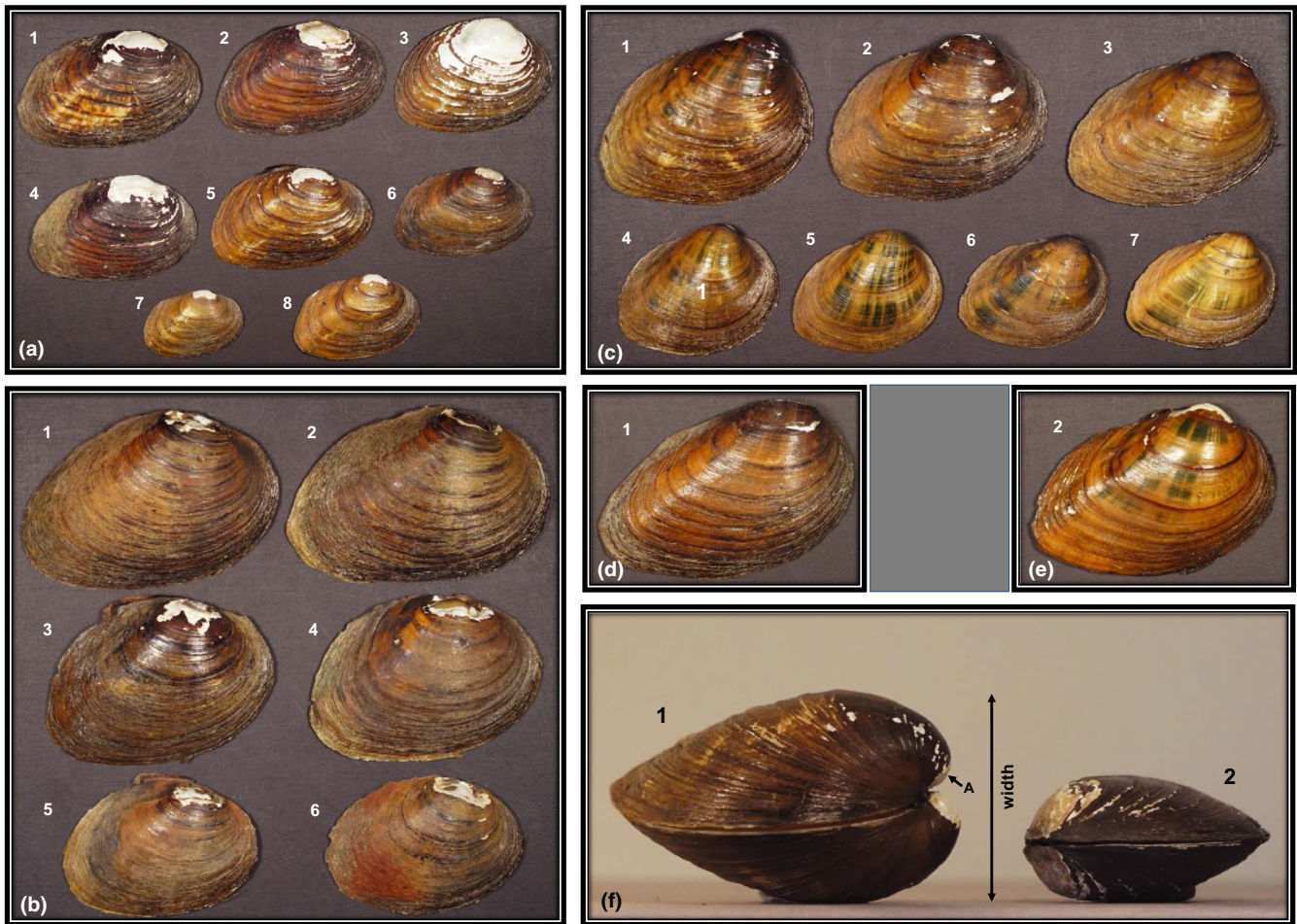
**FIGURE 1** Photographs of individual valves (right only) of *Pleurobema clava* showing morphological variation among shells collected from locations throughout the species range in the Ohio River and Lake Erie watersheds of Ohio (OH), Pennsylvania (PA), and West Virginia (WV), U.S.A. Measurements are for length (L) of individuals in each panel. Panel a: Valves collected from Fish Creek, Maumee River, Lake Erie drainage, Williams County, OH, St. Joseph Township on July 28, 2010, east of CR-18, where (1) L = 48 mm, (2) L = 32 mm, (3) L = 79 mm, and (4) L = 55 mm. McClung Museum Lot # 5143. Panel b: Valves collected from LeBoeuf Creek, Allegheny River drainage, Erie County, PA, on May 12, 1994, east of Route 19, 3 km south of LeBoeuf Gardens where (1) L = 49 mm (2) L = 84 mm, and (3) L = 71 mm. McClung Museum Lot # 0983. Panel c: Valves collected from the Allegheny River, Venango County, PA, President Township on September 26, 1995, at head of island at the mouth of Hemlock Creek, where (1) L = 46 mm, (2) L = 41 mm, (3) L = 38 mm, and (4) L = 63 mm, (5) L = 61 mm. McClung Museum Lot # 16447. Panel d: Valves collected from Elk River, Braxton County, WV, on May 11, 1970, east of Gassaway, 3 miles west of Sutton, Otter Township, where (1) L = 54 mm, and (2) L = 53 mm. McClung Museum Lot # 0978. All photographs taken by J.W. Jones

from distant regions of each species distribution, but in the adjacent regions of their distributions, the morphological similarities and intergradation between the two shell forms remain a source of taxonomic confusion.

The similarities in shell morphologies and disjunct distributions have led some to hypothesize that *P. clava* and *P. oviforme* may encompass either sibling species (USFWS, 1994) or clinal variation of a single species (Ortmann, 1925; Williams et al., 2008), while others have hypothesized that undescribed or cryptic diversity may

exist within the species complex (Haag & Cicerello, 2016; Schilling, 2015). Based on shell morphology, malacologists consider the historical distribution of *P. oviforme* to be confined to the Tennessee and Upper Cumberland River basins, while *P. clava* is widespread throughout the Ohio River basin. *Pleurobema clava* presently occurs in rivers in six states (Michigan, Ohio, Indiana, West Virginia, Kentucky, and Pennsylvania), with large extant populations occurring in the Tippecanoe River, Indiana (Cummings & Berlocher, 1990; Cummings et al., 1992) and in the Allegheny River (Pennsylvania) and





**FIGURE 2** Photographs of individual valves (right only) of *Pleurobema oviforme* showing morphological variation among shells collected from locations throughout the species range in the upper and middle Tennessee River and Cumberland River watersheds of Alabama (AL), Kentucky (KY), Tennessee (TN) and Virginia (VA), U.S.A. Measurements are for length (L) of individuals in each panel. Panel a: Valves collected from the North Fork Holston River, Smyth County, VA, near Riverside on August 8, 1995, where (1) L = 65.5 mm, (2) L = 61.0 mm, (3) L = 59.5 mm, (4) L = 57 mm, (5) L = 58 mm, (6) L = 59.8 mm, (7) L = 45.7, and (8) L = 35.5. McClung Museum Lot # 15462. Panel b: Valves collected from the Hiwassee River, Polk County, TN, at the Route 68 Bridge on July 23, 2007, where (1) L = 53.9 mm, (2) L = 54 mm, (3) L = 48.2 mm, (4) L = 48.3 mm, (5) L = 40.8 mm, and (6) L = 40 mm. McClung Museum Lot #4317. Panel c: Valves collected from the Paint Rock River, Jackson County, AL, at Jones property (RM 34.7) on October 11, 2012, where (1) L = 61.5 mm, (2) L = 63 mm, (3) L = 60.3 mm, (4) L = 46.2 mm, (5) L = 42.3 mm, (6) L = 42 mm, and (7) Length = 41.7. McClung Museum Lot # 5318. Panel d: Valve collected from Buck Creek, Pulaski County, KY, Highway 70 Bridge on August 23, 1999, L = 66.5 mm, McClung Museum Lot #12561. Panel e: Valve collected from Red River, Robertson County, TN, Route 161, 2 miles south of Keysburg on August 27, 1977, L = 47.2 mm, McClung Museum Lot #16432. Panel f: Comparison of valve width of specimens collected from (1) Paint Rock River, L = 62 mm and Width = 33.6 mm (Specimen #1 from B), and (2) North Fork Holston River, Length = 61 mm and Width = 20 mm (Specimen #2 from A). Small arrow A = umbos of specimen 1 protrude past the hinge line, whereas umbos of specimen 2 do not. All photographs taken by J.W. Jones

its tributaries (USFWS, 1994). *Pleurobema oviforme* is considered extant in the middle and upper Tennessee River drainages, including populations in the Clinch, Duck, Hiwassee, North Fork Holston, Paint Rock, Tellico, Pigeon and other rivers, including numerous smaller tributaries to these streams. Additionally, a few small populations are extant in the upper Cumberland River drainage (Haag & Cicerello, 2016; Williams et al., 2008). Although *P. oviforme* remains widespread geographically, the species has lost 58–76% of its former range (Fitzgerald et al., 2021).

Taxonomic assessments have generally recognized *P. clava* and *P. oviforme* as valid species (Turgeon et al., 1998; Williams et al.,

1993, 2017). However, molecular studies utilizing mtDNA have indicated that *P. clava* and *P. oviforme* are closely related (Campbell & Lydeard, 2012; Campbell et al., 2005, 2008), and most recently, molecular evidence for the conspecific status of these species was presented in a comprehensive phylogenetic analysis of the tribe Pleurobemini (Inoue et al., 2018). Unlike *P. clava*, *P. oviforme* is not currently protected under the ESA but is being considered for listing (Federal Register 76[187]:59836–59862); therefore, determining the relationship and resolving taxonomic uncertainties between these taxa has important implications for conservation efforts and decision making regarding listing and protection.

In this study, we analyze morphological and genetic (mitochondrial and nuclear DNA) variation from samples collected throughout the ranges of both *P. clava* and *P. oviforme* to test species boundaries and characterize genetic population structure. First, we assessed genetic distinctiveness by testing monophyly of each putative species expected under the Phylogenetic Species Concept (PSC, Eldridge & Cracraft, 1980; Nelson & Platnick, 1981) using two mitochondrial genes (*COI* and *ND1*). In addition, we used morphometric data collected from museum specimens identified as *P. clava* and *P. oviforme* to test whether the two species can be reliably distinguished based on morphological characters. We then used 13 novel nuclear microsatellite markers to further evaluate species boundaries and determine whether population structuring exists across the broad geographic range of the species complex. We discuss important limitations in our genetic and morphological datasets for delineating these taxa, and additional sampling that is needed to refine our understanding of the two species and potential management units (MUs). These findings have important taxonomic and management implications given the conservation status of *P. clava* and the imminent ESA listing decision for *P. oviforme*.

## 2 | MATERIALS AND METHODS

### 2.1 | Taxon sampling and DNA extraction

Non-lethal tissue samples (mantle clips following Berg et al., 1995 or swabs following Henley et al., 2006) were taken for DNA analyses from individuals of *P. clava* and *P. oviforme* throughout their present ranges in three watersheds (Great Lakes, Ohio, and Tennessee), including 18 drainages in 11 states (Table 1, Figure 3). Due to uncertainty with field identification and differences in sample sizes, individuals were grouped by collection location within drainages (e.g., Allegheny River for *COI* and microsatellites) or by drainage (*ND1*) and will be referred to as "populations" herein (see Table 1 for details). Collection details, museum catalog numbers, and GenBank accession numbers are provided in Morrison et al. (2021).

Genomic DNA was isolated from DNA swabs or mantle tissue preserved in 95% EtOH using the Puregene DNA Extraction Kit (Gentra Systems, Inc.) and resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). DNA concentrations were determined by fluorescence assay (Labarca & Paigen, 1980), and integrity of the DNA was visualized on 1% agarose gels (Sambrook et al., 1989).

### 2.2 | Mitochondrial DNA sequencing and analyses

We sequenced two mitochondrial DNA (mtDNA) protein-encoding genes for phylogenetic and phylogeographic analyses: the *cytochrome oxidase subunit I (COI)* and the *NADH dehydrogenase subunit I (ND1)* genes. Mitochondrial genes were amplified from genomic DNA using the polymerase chain reaction (PCR), with *COI* primers (dgLCO-1490 and dgHCO-2198) and thermal cycling conditions of

Folmer et al. (1994) and *ND1* primers (Leu-urF and LoGlyR) and conditions of Serb et al. (2003). Amplification reactions for both *COI* and *ND1* were performed in 20- $\mu$ l reactions that consisted of the following components and final concentrations: 1 X PCR Buffer (Promega, Madison, WI), 3.75 mM MgCl<sub>2</sub>, 0.3 of each dNTP, 0.2 mM of each primer (from 5 mM stock), 0.5% Tween-20, 0.08 units/ $\mu$ l of *Taq* polymerase (Promega), and 10 ng/ $\mu$ l template DNA. Visual inspection for the targeted amplification product was confirmed on 1% agarose gels.

Polymerase chain reaction products from mtDNA genes were purified with Exonuclease I and Shrimp Alkaline Phosphatase (Promega Corp.) and then used as templates in sequencing reactions with the ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) using the amplification primers. Sequencing reactions were run using a PTC-200 Thermal Cycler (MJ Research) and subjected to electrophoresis on an ABI Prism 3130™ Genetic Analyzer. Products of forward and reverse sequencing reactions were assembled, forming a consensus sequence for each individual and gene using Sequencher 5.4 (GeneCodes Corporation). Multiple alignments were performed for each gene region using Geneious software v. 11.1.5 (<https://www.geneious.com>) with default settings in the Geneious alignment algorithm. The mtDNA alignments were translated into amino acids to ensure the absence of stop codons and indels. Additional sequences (*COI* = 5, *ND1* = 106) were downloaded from GenBank for phylogenetic and phylogeographic analyses (see Morrison et al., 2021 for details).

Population diversity indices, such as the number of segregating sites (*S*), number of haplotypes (*h*), haplotype diversity ( $H_d$ ), and nucleotide diversity ( $\pi$ ), were calculated per population and gene region using DnaSP version 6.12.03x64 (Rozas et al., 2017). Neutrality indices (Tajima's *D* (Tajima, 1989) and Fu's *F<sub>s</sub>* (Fu, 1997)) were calculated for each population and gene region. The average ( $D_{xy}$ ) and net ( $D_n$ ) number of nucleotide substitutions per site were calculated between populations for each gene region using DnaSP. Genetic differentiation among populations was estimated by the haplotype-based method ( $H_s$ ; Hudson et al., 1992), with significance tested via a chi-square table in DnaSP, along with gene flow estimates ( $F_{st}$  and number of migrants,  $N_m$ ; Hudson et al., 1992). Phylogenetic maximum-likelihood (ML) and maximum parsimony (MP) analyses were performed separately for each mtDNA gene. Nucleotide substitution models were determined for six partitions in ModelFinder (Kalyaanamoorthy et al., 2017) for ML phylogenetic analyses: *COI* 1st position—TN+I; *COI* 2nd position—F81+I; *COI* 3rd position—TPM3u+I; *ND1* 1st position—TNe+I; *ND1* 2nd position—HKY+I; and *ND1* 3rd position—TN+I. ML analyses were performed using IQ-TREE v 2.0.6 (Nguyen et al., 2015) with 10 independent runs of initial tree search and 10,000 ultrafast bootstrap replicates (BS) to assess nodal support (Hoang, Chernomor, et al., 2018). MP analyses were performed with default parameters and 1000 ultrafast bootstrap replicates using MPBoot (Hoang, Vinh, et al., 2018). Resulting phylogenies for each mtDNA gene region were reported as 50% majority-rule consensus trees. Haplotype networks were constructed based upon mitochondrial *COI* and *ND1* data separately using the TCS method (Clement et al., 2000) in PopArt v1.7 (Leigh

**TABLE 1** Sampling sites, site codes, and sample sizes (N) for mitochondrial DNA sequences (*COI* and *ND1*) and nuclear microsatellite loci ( $\mu$ sats) for *Pleurobema clava* and *Pleurobema oviforme* included in this study

Taxa	Site	River	State	Drainage	Site code	N <sub>COI</sub>	N <sub>ND1</sub>	N <sub><math>\mu</math>sats</sub>
<i>P. clava</i>	Big Meadow	St. Joseph	MI	Maumee (Lake Erie)	MASJ	13	8	14
<i>P. clava</i>	Fish Creek	Fish Creek	OH	Maumee (Lake Erie)	MAFC	3	4	3
<i>P. clava</i>	West Hickory	Allegheny	PA	Allegheny	ALWH	31	9	33
<i>P. clava</i>	Hunter Station	Allegheny	PA	Allegheny	ALHS	32	10	32
<i>P. clava</i>	Walnut Bend	Allegheny	PA	Allegheny	ALWB	30	6	35
<i>P. clava</i>	Franklin	Allegheny	PA	Allegheny	ALFR	38	1	42
<i>P. clava</i>	French Creek	Allegheny	PA	Allegheny	ALFC	8	2	8
<i>P. clava</i>	Mill Creek	Elk	WV	Elk	ELK	8	2	7
<i>P. clava</i>	Hackers Creek	Hackers Creek	WV	Tygart	TYGT	2	1	5
<i>P. clava</i>	Tippecanoe	Tippecanoe	IN	Wabash	WABA <sup>a</sup>	12	10	17
<i>P. clava</i>	Little Darby	Little Darby	OH	Scioto	SCIO	3	3	3
<i>P. clava</i>	Greensburg	Green	KY	Green	GREN	3	1	5
<i>P. oviforme</i>	Buck/Roundstone Creek	Cumberland	KY	Cumberland	CUMB <sup>a</sup>	7	5	4
<i>P. oviforme</i>	Indian Creek	Clinch	VA	Clinch	CLIN	1	17	1
<i>P. oviforme</i>	Paint Rock	Paint Rock	AL	Paint Rock	PROC	2	3	1
<i>P. oviforme</i>	L. Tennessee <sup>b</sup>	L. Tennessee <sup>b</sup>	NC	L. Tennessee <sup>b</sup>	LTEN <sup>a</sup>	3	13	1
<i>P. oviforme</i>	NF Holston <sup>c</sup>	NF Holston <sup>c</sup>	VA	Holston	HOLS	28	59	24
<i>P. oviforme</i>	Lillard Mill	Duck	TN	Duck	DUCK	0	3	0
<i>P. oviforme</i>	Sevierville	Little Pigeon	TN	French Broad	FRBR	0	3	0
<i>P. oviforme</i>	Turtletown	Hiwassee	TN	Hiwassee	HIWA	0	5	0
<i>P. oviforme</i>	Coulter's Bridge	Little River	TN	Tennessee	TENN	0	2	0
<i>P. oviforme</i>	Sinking Spring	Little Chucky	TN	Nolichucky	NOLI	0	3	0
<i>P. oviforme</i>	Ringgold <sup>a</sup>	S Chickamauga	GA	S Chickamauga	SCHK <sup>a</sup>	0	8	0
Total						224	178	235

Note: For microsatellite analyses, the single samples from the Clinch, Paint Rock, and Little Tennessee were combined as they all originated from the Tennessee watershed (TENN).

<sup>a</sup>Includes multiple sampling sites, see Morrison et al. (2021) for details.

<sup>b</sup>Little Tennessee.

<sup>c</sup>North Fork Holston.

& Bryant, 2015), which follows the parsimony-based approach of Templeton et al. (1992, "TCS" method, after the authors).

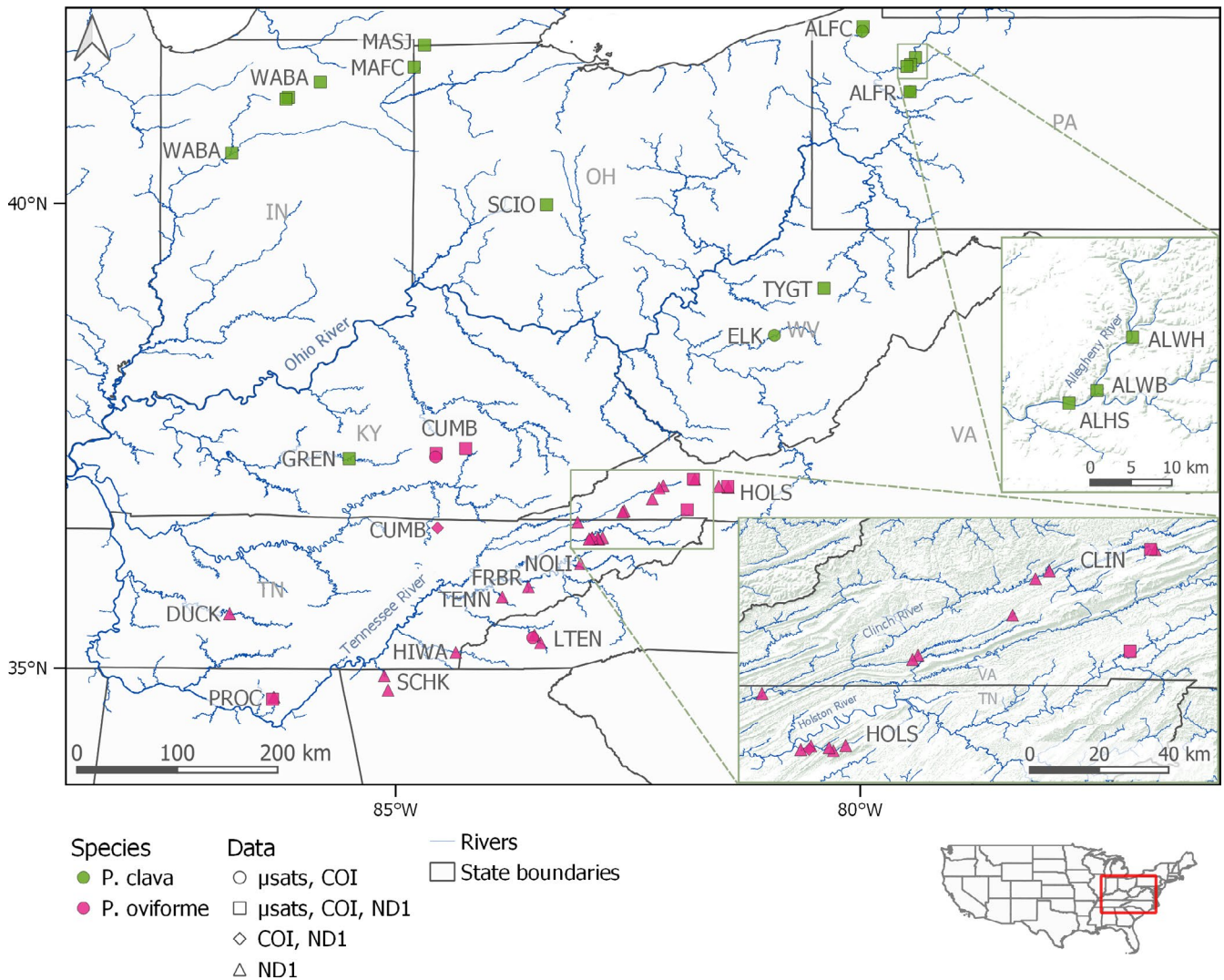
### 2.3 | Morphometrics

We used museum specimens (McClung Museum of Natural History and Culture—Knoxville, TN, and Florida Museum of Natural History—Gainesville, FL) identified as *P. clava* and *P. oviforme* to evaluate morphological variation throughout the range of both species. A single valve from each individual was positioned with ventral margins of anterior and posterior adductor muscle scars on a horizontal plane before taking the measurements to the nearest 0.01 mm using digital calipers. Total length anterior to posterior (length) parallel to the adductor scar plane, width at interdentum (width), and height from intersection of hinge ligament and umbo to the ventral margin of

the shell (height) parallel to the adductor scar plane were measured. Valves were weighed to the nearest 0.1 g using a digital scale.

We used  $\log_e$ -transformed measurement values to produce a scale-invariant matrix while preserving information about allometry (Kowalewski et al., 1997; Strauss, 1985). We used the three  $\log_e$ -transformed variables to calculate three ratios: height/length, width/length, and width/height. We used principal components analysis (PCA) and random forest analysis (RF) applied to standardized ratios to examine morphological variation among species. The random forest analysis was conducted using 80% (238 individuals) of the data as the training dataset and 20% (60 individuals) of the data as the validating dataset. After plotting the out-of-bag (OOB) error rate and observing how it declined and stabilized, a total of 500 trees was used for the analysis. The random forest analysis was conducted using the "rattle" package (Williams, 2011) as implemented in program R (R Core Team, 2019). To further evaluate the ability of





**FIGURE 3** Sampling locations for *Pleurobema clava* (green) and *Pleurobema oviforme* (red) samples analyzed in this study, with shapes indicating inclusion in different datasets. Site codes are given in Table 1

morphological measurements to correctly classify individual shells into each respective species category, we conducted a linear discriminant analysis (LDA) with cross-validation on data rarified to the lowest sample size ( $N = 129$ ) using the *lda* function in the MASS package (Venables & Ripley, 2002) in R v. 4.0.2 (R Core Team, 2019).

## 2.4 | Microsatellite genotyping and population genetic analyses

Two microsatellite-enriched libraries were prepared by Genetic Identification Services, Inc. (GIS, <http://www.genetic-id-services.com/>), using magnetic bead capture technology (Peacock et al., 2002) and tetranucleotide microsatellite motif repeat capture molecules TACA and TAGA. Three hundred clones that contained repeats of appropriate length (>10 repeats) and had adequate flanking regions were targeted for primer development, including a modified 19-bp M13 tag on the 5' end of each forward primer (Boutin-Ganache et al.,

2001). PCR conditions were optimized as described in King et al. (2006), using genomic DNA from four individuals of *P. clava* from the Allegheny French Creek (ALFC), Maumee St. Joseph (MASJ), and Fish Creek (MAFC) populations plus four *P. oviforme* from the Cumberland drainage.

Based on size, strength of PCR amplification, ease of scoring, and observed heterozygosity, 13 loci were labeled utilizing the ABI Prism™ 5-dye filter set (FAM, VIC, NED, PET, and LIZ). Loci were amplified individually, and PCR products were multiplexed into four size- and dye-compatible groups. Microsatellite DNA amplification reactions consisted of 1 X PCR Buffer (Promega), 3.75 mM MgCl<sub>2</sub>, 0.3 of each dNTP, 0.2 mM of each primer (from 5 mM stock), 0.5% Tween-20, 0.08 units/μl of *Taq* polymerase (Promega), and 20 ng/μl template DNA in a volume of 10 μl. Amplifications were carried out using the following thermal cycler profile: initial denaturing at 94°C for 2 min; 35 cycles of 94°C C for 40 s, 58°C C for 40 sec, 72°C C for 1 min; and a final extension at 72°C C for 5 min. Amplified, labeled PCR products were subjected to capillary electrophoresis on

an ABI Prism™ 3130 XL Genetic Analyzer (Applied Biosystems, Inc.) with Gene Scan 500 LIZ size standards. GeneScan 3.7 analysis software, Genotyper 4.0 (Applied Biosystems), and GeneMapper (ver. 4) (Thermo Fisher Scientific) were used to score, bin, and output allelic (and genotypic) data following the protocols described in King et al. (2001).

The quality of microsatellite genotyping was checked by several measures. First, duplicate multi-locus genotypes (MLGs) were located using GenAEx v. 6.5 (Peakall & Smouse, 2006). MICROCHECKER v. 2.2 (Van Oosterhout et al., 2004) was used to check for genotyping errors, large allele dropout, and segregation of null alleles. Fisher's exact test was used to test for linkage disequilibrium (LD), and exact tests for Hardy–Weinberg equilibrium (HWE) were performed in GENEPOP on the Web (Raymond & Rousset, 1995). Sequential Bonferroni adjustments were applied for multiple tests (Rice, 1989). The power to distinguish unique MLGs was calculated as the probability of identity ( $P_{ID}$ ; Peakall & Smouse, 2006) in GenAEx. Since sample sizes varied among populations, allelic richness, or the number of alleles per population, was calculated by rarefaction (Kalinowski, 2004) to compensate for unequal sample sizes using HP-Rare (Kalinowski, 2005). Kinship coefficients were calculated among all individuals in Genodive v. 3.04 (Meirmans, 2020).

Traditional *F*-statistics (Wright, 1951) were used to test for population structuring within and among drainages and watersheds for microsatellite data using analysis of molecular variance (AMOVA, Excoffier et al., 1992; Michalakis & Excoffier, 1996) in GenAEx. To examine pairwise population structuring, we calculated two complementary measures in Genodive: the fixation index  $G'_{st}$  (Nei, 1987) and Jost's *D* (Jost, 2008), assessing significance with 9,999 permutations. Relationships among populations were assessed via the neighbor-joining method (Saitou & Nei, 1987) using Nei's standard genetic distances ( $D_A$ , Nei, 1972) in Poptree2 (Takezaki et al., 2010). Population-based microsatellite analyses included grouping of single samples from the Little Tennessee (LTEN), Paint Rock (PROC), and Clinch (CLIN) collections into a Tennessee (TENN) population (Table 1).

Two different clustering methods were utilized to explore potential structuring within the MLG dataset. First, a Bayesian model-based clustering approach was implemented in STRUCTURE v. 2.4 (Pritchard et al., 2000) to determine the number of clusters (*K*) within the MLG dataset by minimizing deviation from HWE within clusters. Values of *K* = 1 to 15 were tested, where *K* = 1 represents a single panmictic population and *K* = 15 represents the 15 populations that were sampled in this study. Twenty replicate simulations were run with 200,000 Markov chain Monte Carlo repetitions and a burn-in of 100,000 with an admixture model, assuming independent allele frequencies across populations as priors (Hubitz et al., 2009). Results from replicate iterations were summarized using default settings in CLUMPAK (Kopelman et al., 2015) implemented in StructureSelector (Li & Liu, 2018); where the Evanno method ( $\Delta K$ , Evanno et al., 2005) and the posterior probability (PP) of each *K*-value across replicates were calculated following Bayes' rule (p. 13, Pritchard & Wen, 2004) noted as  $\ln \Pr(X|K)$  (Pritchard et al., 2000) and were used to discern

the best-supported number of clusters using both the maximum-likelihood scores and  $\Delta K$  methods. Second, a discriminant analysis of principal components (DAPC; Jombart et al., 2010) analysis was performed in the R package adegenet (Jombart, 2008). Unlike the STRUCTURE analysis, the DAPC analysis clusters genetically similar individuals in multivariate space based on allelic composition without reliance on HWE and LD assumptions. Genetic isolation by distance (IBD) was analyzed using a Mantel test (Mantel, 1967) comparing pairwise genetic distances among populations (Nei's  $D_A$ ) with geographic river distances in adegenet. River distances were calculated as the shortest path between populations along U.S. National Hydrography flowlines (NHDPlus version 2) referenced to the Albers Equal Area North American Datum 1983 Coordinate Reference System (EPSG = 42303) using the packages sf and stplanr (Lovelace & Ellison, 2018; Pebesma, 2018; R Core Team, 2019).

The program BOTTLENECK v 1.2.02 (Cornuet & Luikart, 1996) was used to determine whether any populations experienced a recent reduction in population size, or population bottleneck, which would produce evidence of heterozygote excess (*H<sub>x</sub>*; Cornuet & Luikart, 1996; Luikart et al., 1998). Two mutation models were calculated, including the conservative stepwise mutation model (SMM) and the two-phase model (TPM) allowing for both stepwise and multi-step mutations, as assumed to be reasonable for most microsatellites (Di Rienzo et al., 1994). For the TPM model, a variance of 12 was used (Piry et al., 1999) and the proportion of stepwise mutations in the model was set to 90 (Dussex et al., 2015; Garza & Williamson, 2001). Significance of heterozygote excess was tested using the sign test (Cornuet & Luikart, 1996), the Wilcoxon signed-rank test (Luikart & Cornuet, 1998), and the mode-shift indicator (Luikart et al., 1998) based on the presence of a mode shift in the distribution of allele frequencies.

### 3 | RESULTS

#### 3.1 | Mitochondrial DNA

We generated 219 *COI* and 70 *ND1* DNA sequences for this study (Table 1). All novel sequences generated from this study are accessible on GenBank (*COI*: MT991776–MT991982; *ND1*: MW005982–MW006051). The *P. clava/oviforme* *COI* sequence alignment of 224 sequences was 572 bp in length, including 29 variable sites and nucleotide diversity  $\pi = 0.0061$ . The *ND1* alignment of 176 sequences was trimmed to 775 bp in length to minimize missing data, contained 55 variable sites, and nucleotide diversity  $\pi = 0.0049$ . No gaps or stop codons were detected in either protein-encoding gene sequence.

Summary statistics for nucleotide and haplotype diversity for the *COI* and *ND1* *P. clava/oviforme* datasets are shown in Table 2. The number of segregating sites (*S*) was higher for *ND1* (55) than for *COI* (29). Similarly, haplotype diversity ( $H_d$ ), which accounts for different sample sizes, was higher for the *ND1* compared with *COI* datasets (0.933 vs. 0.564, respectively). Haplotype diversity ( $H_d$ )



**TABLE 2** Summary of DNA sequence variation at the mitochondrial *COI* and *ND1* genes for *Pleurobema clava* and *Pleurobema oviforme*, where  $n$  is the number of sequences,  $S$  is the number of segregating sites,  $h$  is the number of haplotypes (number of haplotypes unique to population),  $H_d$  is haplotype diversity,  $k$  is the number of pairwise differences, and  $\pi$  is the nucleotide diversity

Gene	Taxon	Site	$n$	$S$	$h$	$H_d$	$k$	$\pi$
COI (572 bp)	<i>P. clava</i>	MASJ	13	9	3 (0)	0.4103	2.4103	0.0042
	<i>P. clava</i>	MAFC	3	3	2 (0)	0.6667	2	0.0035
	<i>P. clava</i>	ALHS	32	9	3 (0)	0.2319	1.869	0.0033
	<i>P. clava</i>	ALWH	31	10	4 (0)	0.4323	1.8495	0.0033
	<i>P. clava</i>	ALWB	30	6	3 (0)	0.131	0.4	0.0007
	<i>P. clava</i>	ALFR	38	15	4 (1)	0.2006	0.8393	0.0015
	<i>P. clava</i>	ALFC	8	9	2 (1)	0.25	2.25	0.004
	<i>P. clava</i>	ELK	8	9	2 (0)	0.25	2.25	0.004
	<i>P. clava</i>	TYGT	2	0	1 (0)	0	0	0
	<i>P. clava</i>	WABA <sup>1</sup>	12	10	3 (1)	0.3182	1.6667	0.0029
	<i>P. clava</i>	SCIO	3	0	1 (0)	0	0	0
	<i>P. clava</i>	GREN	3	2	2 (0)	0.6667	1.3333	0.0023
	<i>P. oviforme</i>	CUMB <sup>1</sup>	7	2	2 (1)	0.476	0.952	0.002
	<i>P. oviforme</i>	HOLS	28	15	8 (5)	0.8228	4.3942	0.0077
	<i>P. oviforme</i>	LTEN	3	5	2 (2)	0.667	3.333	0.0058
	<i>P. oviforme</i>	PROC	2	4	2 (1)	1	4.000	0.0070
	<i>P. oviforme</i>	CLIN	1	NA	1 (1)	NA	NA	NA
Average						0.564	3.448	0.0061
Total			224	29	20 (13)			
ND1 (775 bp)	<i>P. clava</i>	MASJ	8	6	3 (0)	0.4643	1.8571	0.0027
	<i>P. clava</i>	MAFC	4	5	3 (0)	0.8333	3.1667	0.0046
	<i>P. clava</i>	ALLE <sup>1</sup>	28	11	8 (2)	0.6217	1.7355	0.0025
	<i>P. clava</i>	ELK	2	6	2 (0)	1	6	0.0077
	<i>P. clava</i>	WABA <sup>1</sup>	10	8	6 (1)	0.9111	2.8889	0.0042
	<i>P. clava</i>	TYGT	1	NA	1 (0)	NA	NA	NA
	<i>P. clava</i>	SCIO	3	3	3 (0)	1	2	0.0028
	<i>P. clava</i>	GREN	1	NA	1 (0)	NA	NA	NA
	<i>P. oviforme</i>	CUMB <sup>1</sup>	5	7	3 (3)	0.8333	3.5	0.0051
	<i>P. oviforme</i>	CLIN <sup>1</sup>	17	12	8 (6)	0.8529	2.7353	0.004
	<i>P. oviforme</i>	PROC	3	4	2 (1)	0.6667	2.6667	0.0039
	<i>P. oviforme</i>	HOLS <sup>1</sup>	59	18	15 (7)	0.886	2.9515	0.0043
	<i>P. oviforme</i>	LTEN	13	6	3 (1)	0.7	2.8	0.0041
	<i>P. oviforme</i>	DUCK	3	0	1 (0)	0	0	0
	<i>P. oviforme</i>	FRBR	3	7	3 (1)	1	4.6667	0.0068
	<i>P. oviforme</i>	HIWA	5	8	4 (3)	0.9	4.4	0.0064
	<i>P. oviforme</i>	TENN	2	4	2 (2)	1	4	0.0059
	<i>P. oviforme</i>	NOLI	3	6	3 (1)	1	4	0.0059
	<i>P. oviforme</i>	SCHK <sup>1</sup>	8	6	6 (4)	0.8	2.4667	0.0036
	Average						0.933	3.364
Total			178	55	45 (32)			

Note: Site refers to site codes given in Table 1. Note that diversity statistics for the CLIN (*COI*) and TYGT/GREN populations (*ND1*) are omitted because only one sequence was available.

<sup>a</sup>Includes multiple sampling sites, see Table 1 and Morrison et al. (2021).

varied substantially among populations for COI, ranging from 0 in the Tygart (TYGT), Scioto (SCIO), and LTEN populations to 0.8228 in the HOLS and was consistently low in the five Allegheny (ALLE) and Elk (ELK; Ohio drainage) populations (0.131–0.25; Table 2), whereas  $H_d$  was high for most populations at ND1 (0.62–1), with the exception of St. Joseph/Maumee ([MASJ], 0.4103; Table 2).

For the COI dataset, uncorrected sequence divergence ( $D_{xy}$ ) between populations ranged from 0 to 1.69 (Table S1). Sequence divergence ( $D_{xy}$ ) among *P. clava* populations averaged 0.55% (range = 0–1.34%), with the highest values involving comparisons with the Maumee drainage (range = 1.08–1.34%). Interestingly, the *P. clava* Maumee populations were more closely related to *P. oviforme* populations from the Cumberland (CUMB), PROC, HOLS, and LTEN populations ( $D_{xy}$  range = 0.32–0.66%) than to more proximate *P. clava* populations. Divergence was low for populations within the Allegheny drainage, including the TYGT (0.11–0.34%), as well as among the Wabash (WABA), Green (GREN), and SCIO populations (0–0.22%). Differentiation was greater among *P. oviforme* populations (average  $D_{xy}$  = 0.88%, range 0.38–1.57%), except for lower divergence between the LTEN and CUMB populations (0.38%). Between-species comparisons produced a slightly higher average than within *P. oviforme* populations (0.99%), but a range of divergences similar to that observed between *P. oviforme* populations (range = 0.32–1.69%). The highest pairwise comparisons involved *P. clava* versus *P. oviforme* from the PROC drainage ( $D_{xy}$  range = 1.43–1.69%).

Similar patterns were observed when net nucleotide differences for the COI dataset were examined (Table S1), with the Maumee drainage producing the highest sequence divergence among *P. clava* populations ( $D_a$  range = 0.69–1.05%) yet lower  $D_a$  in comparisons with *P. oviforme* from the CUMB, PROC, and HOLS populations ( $D_a$  range = –0.09–0.23%). No to minimal net nucleotide substitutions were observed among remaining *P. clava* populations. Greatest  $D_a$  distances were observed in comparisons with the PROC drainage ( $D_a$  range = 0.90–1.22%, except for the Maumee and CUMB populations mentioned above). Among all collections, genetic differentiation was significant among populations, with  $H_s$  = 0.350 ( $p$  < .001). Gene flow estimates indicated an  $F_{st}$  = 0.571 and number of migrants ( $N_m$ ) = 0.19. No significant deviations from neutrality were detected by Tajima's  $D$  = –0.936 ( $p$  > .10) or Fu's  $F_s$  = –2.310 ( $p$  > .05).

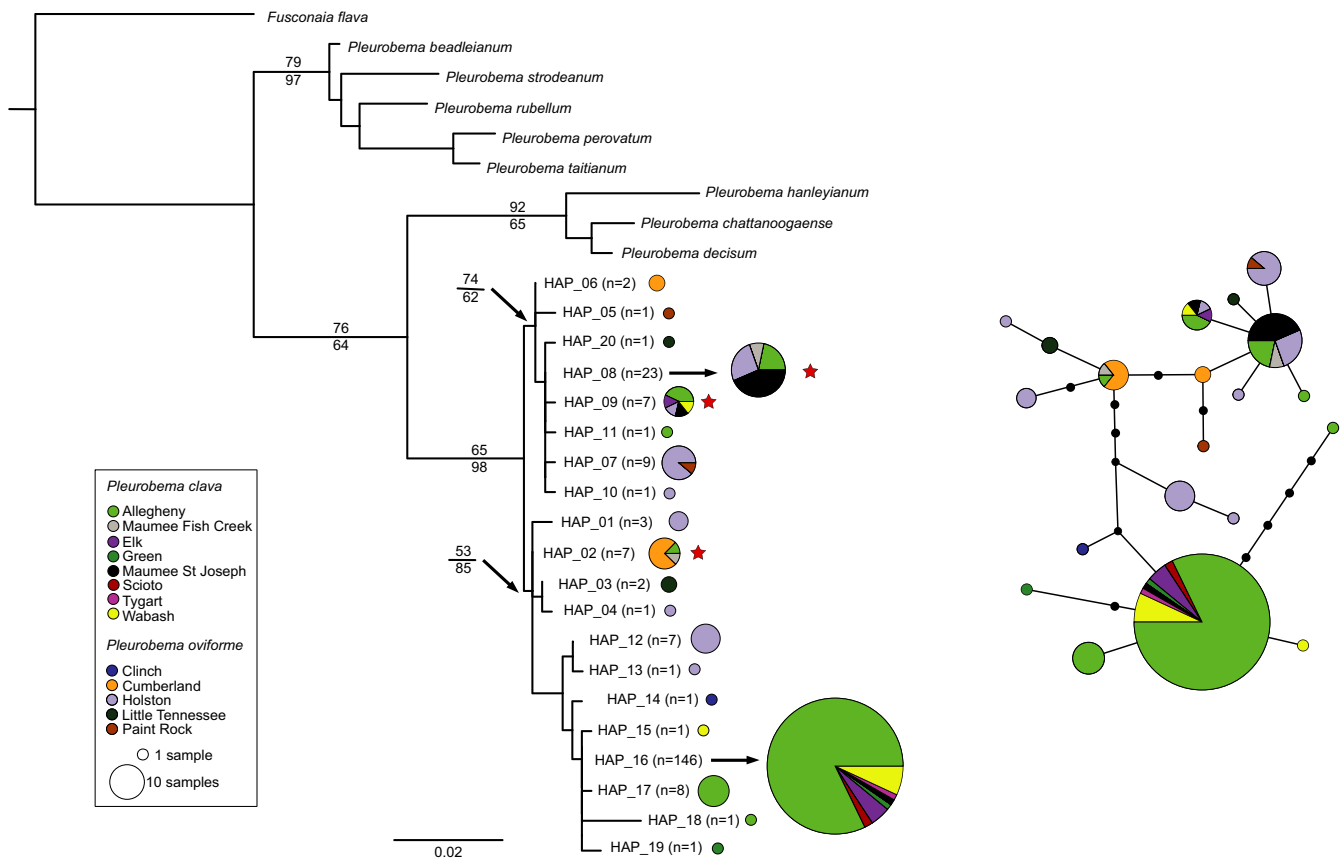
Uncorrected pairwise percent sequence divergence ( $D_{xy}$ ) at ND1 ranged from 0 (*P. clava* from the GREN and TYGT) to 1.2% (*P. oviforme* from the PROC and Tennessee (TENN); Table S2). Among *P. clava* populations, sequence divergence averaged 0.47% (range = 0.14–0.85%) while comparisons among *P. oviforme* populations were slightly higher, averaging 0.77% (range = 0.34–1.20%). Like the COI analyses, the highest divergence estimates among *P. clava* populations involved the Maumee populations (0.45–0.85%). It is noteworthy that the MASJ population was more closely related to the *P. oviforme* Duck (DUCK) and South Chickamauga (SCHK) populations (~0.4%) than to other *P. clava* populations (0.52–0.85%). Among *P. oviforme* populations, the PROC/DUCK and the DUCK/SCHK comparisons were lowest ( $D_{xy}$  = 0.34%). Pairwise comparisons between putative

species were intermediate, ranging from 0.26% (TENN/TYGT) to 1.05% (TENN/MASJ), and averaging 0.66%. Minimal net nucleotide substitutions ( $D_a$ ) were detected between several populations, even those involving the two species (e.g., *P. oviforme* from the LTEN, CLIN, and CUMB compared with *P. clava* from the ELK, TYGT, GREN, and SCIO populations). Like the COI dataset, genetic differentiation was significant among all collections, with  $H_s$  = 0.7930 ( $p$  < .001); however, the fixation index was lower and gene flow estimates were slightly higher for ND1 ( $F_{st}$  = 0.2459,  $N_m$  = 0.77, respectively). While the COI dataset followed expectations under neutrality, the ND1 dataset deviated, with significant values for both Tajima's  $D$  (–1.964,  $p$  < .05) and Fu's  $F_s$  (–34.060,  $p$  < .001), indicating an excess of rare alleles (Tajima, 1989), suggestive of either a rapid expansion from a population with small effective size (population bottleneck) or positive selection (a selective sweep).

Maximum-likelihood (ML) and maximum parsimony (MP) phylogenetic analyses included 29 COI sequences, of which 20 were the unique *P. clava* and *P. oviforme* haplotypes from this study, plus nine sequences from GenBank representing several closely related *Pleurobema* species (i.e., *P. decisum*, *P. hanleyanum*, *P. beadelianum*, *P. rubellum*; Campbell et al., 2005; Inoue et al., 2018; Figure 4a). The *P. clava* and *P. oviforme* sequences formed a single clade that was better supported by MP than ML (support values MP = 98, ML = 65; Figure 4a). Although structuring among *P. clava* and *P. oviforme* haplotypes was weak, two clades were evident, each containing both shared and species-specific haplotypes. The first clade (support values ML = 74 and MP = 62; Figure 4a) included five *P. oviforme*-specific (5, 6, 7, 10, and 20) and one *P. clava*-specific (11) haplotypes as well as two of the three shared haplotypes (8, 9). The second clade (support values ML = 53 and MP = 85; Figure 4a) contained the most common haplotype (16; Figure 4a) that included most of the *P. clava* sequences, plus four additional *P. clava*-specific (15, 17, 18, and 19) haplotypes, as well as six *P. oviforme*-specific (1, 3, 4, 12, 13, and 14) and one shared haplotype (2; Table S3).

The two clades recovered in the ML and MP analyses were apparent in the COI TCS haplotype network (Figure 4b). It was noteworthy that haplotypes differed by few (1 to 4) substitutions. Shared haplotypes were either intermediate in the network (e.g., Haplotype 2; Figure 4b) or at a distal portion of the network corresponding to Clade 1 in the ML analysis (Figure 4a). Most of the *P. clava* samples had Haplotype 16 ( $N$  = 146; Figure 4b, Table S3), which was common in the ALLE populations, but was also found in geographically distant populations such as the WABA, MASJ, Maumee Fish Creek (MAFC), ELK, TYGT, and GREN populations. Intermediate haplotypes were comprised mostly of *P. oviforme* sequences from the HOLS and CUMB populations.

A greater representation of *P. oviforme* was achieved in the ND1 datasets after the addition of sequence data representing 11 populations from Schilling (2015; Table 1). Maximum-likelihood and MP phylogenetic analyses included 60 ND1 sequences, including 45 unique *P. clava* and *P. oviforme* haplotypes, an undescribed but closely related taxon (*P. cf. oviforme*, Schilling, 2015), plus several



**FIGURE 4** Mitochondrial *COI* (a) maximum-likelihood (ML) phylogeny (bootstrap support values for ML above and maximum parsimony below branches) and (b) TCS haplotype network for *Pleurobema clava* and *Pleurobema oviforme*. Red stars designate haplotypes present in both species. Colors within pie diagrams indicate populations where given haplotypes were observed

related *Pleurobema* species. As reported in the analysis conducted by Schilling (2015), *P. cf. oviforme* was the sister group to *P. clava/oviforme* (Figure 5a; uncorrected *p*-distance = 5.5%, not shown). As in the *COI* phylogenetic analysis, *P. clava/oviforme* haplotypes formed a single, well-supported clade (support values of 100 for both ML and MP; Figure 5a) with minimal differentiation within.

The *ND1* haplotype network incorporating 178 sequences and 45 unique haplotypes formed a star-like network containing two common central haplotypes with many closely related haplotypes (differing by few substitutions) surrounding the common haplotypes (Figure 5b). Thirty-four and eight haplotypes were unique to *P. oviforme* and *P. clava*, respectively, while four haplotypes were shared between species (Figure 5b, Table S4), and similarly to observations of the *COI* network (Figure 4b), haplotypes differed by few (1–4) substitutions. The two most common haplotypes (Haplotypes 4 and 35; Figure 5b, Table S4) differed by a single substitution. Haplotype 4 was widely distributed geographically and common in both species, while Haplotype 35 was comprised of *P. oviforme* from the HOLS and LTEN. Three additional shared haplotypes (17, 29, and 36) differed from the most common haplotypes by one to two substitutions. Haplotypes from the HOLS were intermingled throughout the network, along with species-specific haplotypes.

### 3.2 | Morphometrics

Morphometric measurements were taken from 363 museum specimens (*P. clava* *N* = 129; *P. oviforme* *N* = 234). The first two PCs described 86.7% of the total variation (Figure 6). The PCA showed broad overlap in morphometric measurements between species (Figure 6a), with all variables besides the length/height ratio loading equally on PC1. PC2 largely described a gradient in length (total length, length/height, length/width). Shells collected from the upper and middle Tennessee drainage were separated on the upper and lower quadrants of the PCA plot (Figure 6b). The PERMANOVA identified significant differences between the nominal species; however, these differences explained a small percentage of the total morphological variation present ( $R^2 = 0.030$ ,  $F = 11.344$ ,  $p < .001$ ). Classification to species by the LDA was 58.9% and 72.1% accuracy for *P. clava* and *P. oviforme*, respectively, with an overall classification accuracy of 65.5% (Table 3).

For the RF analyses, the number of variables tested at each split was set to 2 based on the out-of-bag (OOB) error rate estimate. The OOB error rate was 16.8%, indicating that the model had approximately an 83% accuracy for identification of shells to the correct species (Table 3). Based on the mean decrease in the Gini scores, the



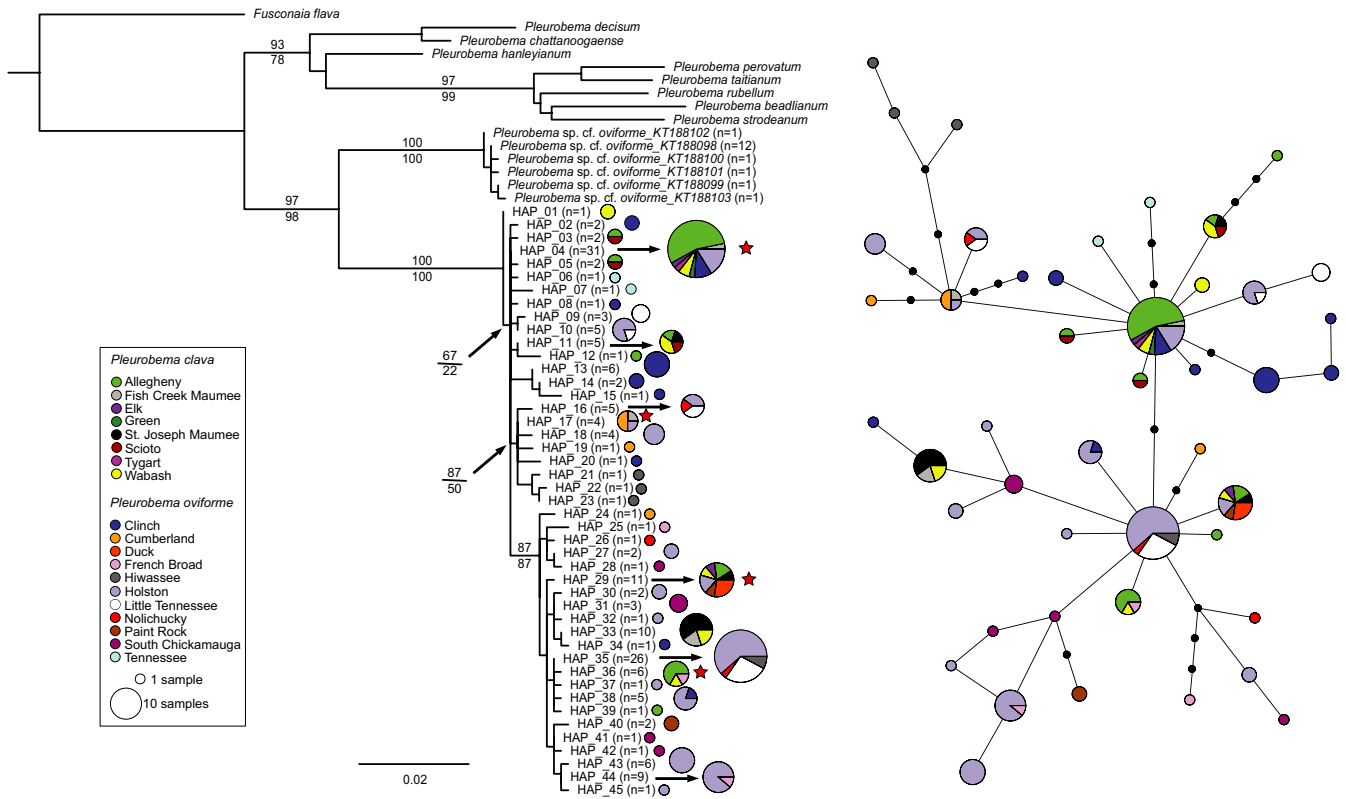


FIGURE 5 Mitochondrial ND1 (a) maximum-likelihood phylogeny (bootstrap support values for ML above and maximum parsimony below branches) and (b) TCS haplotype network for *Pleurobema clava* and *Pleurobema oviforme*. Red stars designate haplotypes present in both species. Colors within pie diagrams indicate populations where given haplotypes were observed

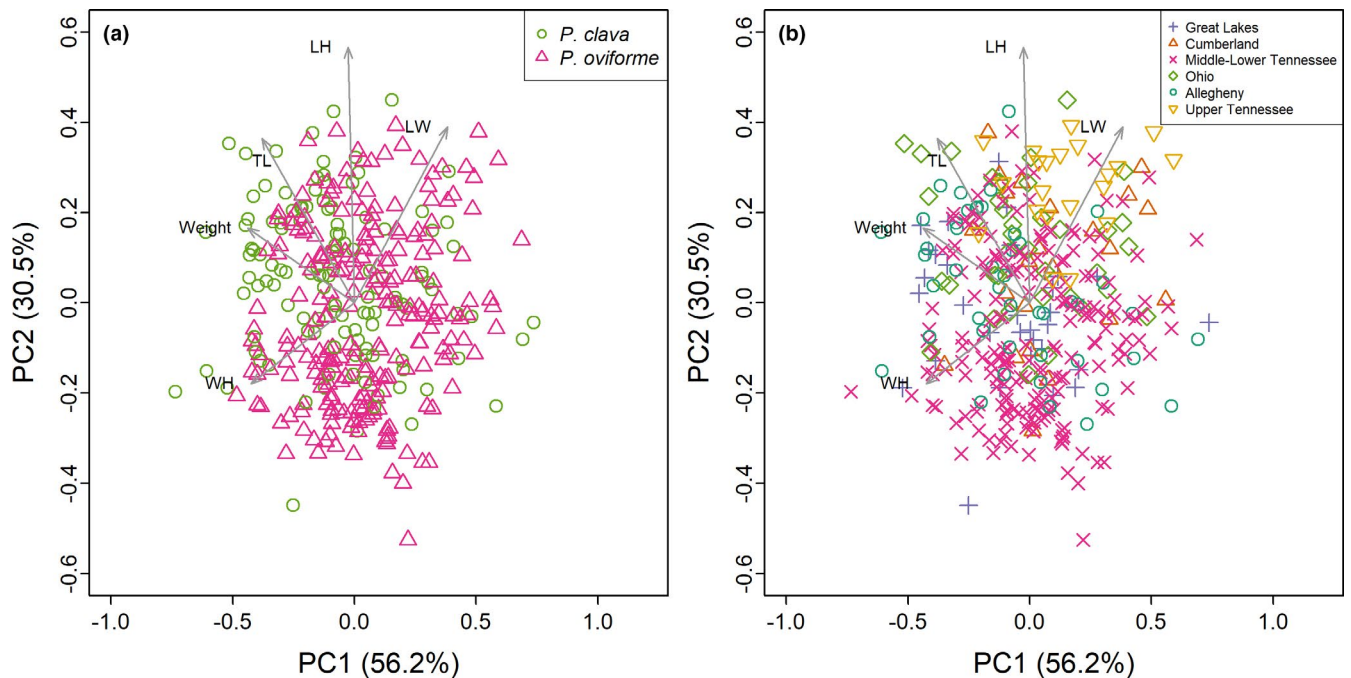


FIGURE 6 Principal Coordinates Analysis of morphometric data with arrows for log-transformed biplot variables (WH = width/height; TL = total length; LH = length/height; LW = length/width), with data colored by putative species (a) and by region of collection (b)

**TABLE 3** Results of morphological analyses of four shell ratio variables (width to height, width to length, height to length, and weight to length) taken from museum samples of *Pleurobema clava* and *Pleurobema oviforme*

A. Correct Class	<i>P. clava</i>	<i>P. oviforme</i>	Percent correct classification
<i>P. clava</i>	76	53	58.9%
<i>P. oviforme</i>	36	93	72.1%
Overall			65.5%
B. Correct Class	<i>P. clava</i>	<i>P. oviforme</i>	Percent correct classification
<i>P. clava</i>	112	23	83.0%
<i>P. oviforme</i>	27	136	83.4%
Overall			83.2%

Note: Each confusion matrix shows the number of correctly and incorrectly identified individuals for each species. A. Cross-validation linear discriminant analysis. B. Random forest analysis.

ratio of shell width to shell height was the most important variable for classification of shells to the correct species, followed by width to length, height to length, and weight to length.

### 3.3 | Nuclear microsatellite markers

Of 300 primer pairs tested for amplification, 13 loci were polymorphic and consistently produced alleles within expected size ranges (Table S5). Although several pairwise comparisons for linkage disequilibrium remained significant after correction for multiple tests (adjusted  $\alpha = 0.0006$ ), none were consistent across populations (results not shown). Similarly, only one in 388 tests for departure from HWE was significant (*Pcl278* in the VA\_NFH population; Table S6). MICROCHECKER did not detect evidence for large allele dropout or scoring errors; however, null alleles were detected in 10 of 130 locus-by-population tests, yet none were consistent across populations. Two MLGs (*Pcla\_PA\_AR\_HS\_02* and *Pcla\_PA\_AR\_HS\_13*) differed at only one locus (*PclC211*) and had a kinship value of 0.443. Given the low probability that two unrelated individuals sampled from a single population share the same MLG ( $PI = 4.0 \times 10^{-14}$ ;  $PI_{sibs} = 2.6 \times 10^{-6}$ ), this high kinship estimate suggests a possible parent-offspring relationship.

A total of 454 alleles were detected among 235 genotyped samples (Table S5). The number of alleles per locus ranged from 23 (*PclD7*) to 63 (*PclC211*) and averaged 11.5 (Table S6). The mean number of alleles per population ranged from 4.39 (SCIO) to 21 (ALLE Franklin (ALFR)), with a mean of 11.5 (Table S6). Following rarefaction to compensate for unequal sample sizes, the WABA and several ALLE populations had the highest allelic richness (5.05–5.15) while the HOLS had the lowest (4.24). Private alleles were detected within all populations, with the MASJ and HOLS populations having most (14 and 17, respectively), while a single private allele was observed in the ELK and GREN populations. Similarly, rarefied private allelic richness estimates were higher for the MASJ and HOLS populations

(0.87 and 0.84, respectively), but the TENN and WABA also had high values (0.92 and 0.80, respectively). Observed levels of heterozygosity were generally high, ranging from 0.583 (CUMB) to 0.915 (ALFR), with a global mean of 0.829. Most inbreeding coefficients were negative or close to 0, but two of the *P. oviforme* populations (CUMB and TENN) plus the *P. clava* GREN population had positive inbreeding coefficients, suggesting heterozygote deficits. A likely explanation for these positive inbreeding coefficients is a Wahlund effect given that these samples originated from more than one location within the drainage.

We found little evidence for recent bottlenecks using the microsatellite data from the seven populations with sample sizes of ten or more individuals. No significant results were obtained using the heterozygote excess method (*Hx*) for either the SMM or TPM mutation models ( $p > .00156$ ; Tables S7 and S8), and no significant deviations from a normal L-shaped distribution were detected (mode shift). Two of the ALLE populations showed significant heterozygote deficits ( $H_d$ ) under one (ALLE Hunter Station (ALHS)) or both (ALLE West Hickory (ALWH)) mutation models, suggesting that these populations are not in mutation-drift equilibrium (Table S8).

Several biological hypotheses were assessed using fixation measures ( $F_{st}$ ) and AMOVA. First, we tested whether genetic structuring was apparent among all 15 populations. An AMOVA with populations as a second hierarchical level produced a weak but significant global  $F_{st}$  value of 0.029, with 3% of variation among populations and 97% within populations (Table S9A). We next tested whether greater structuring was apparent between the two putative species, and a significant and greater global  $F_{st}$  resulted (global  $F_{st} = 0.059$ ,  $p \leq .0001$ , 4% of variation among putative species; Table S9B). Given that genetic structuring often occurs between watersheds in other unionid species (Galbraith et al., 2015), we also tested whether significant structuring existed between the three watersheds represented in our dataset. Significant structuring was detected, yet the magnitude of differentiation was lower than between species (global  $F_{st} = 0.049$ ,  $p \leq .0001$ , 3% of variation among watersheds; Table S9C).

Pairwise fixation indices ( $G'_{ST}$ ) among *P. clava* populations ranged from 0 (no genetic differentiation, ALLE, MAFC/WABA, and ALLE Walnut Bend (ALWB)/TYGT populations) to 0.056 (moderate genetic differentiation, MASJ vs. SCIO; Table 4) and averaged 0.022 (little genetic differentiation; Hartl & Clark, 1997). All pairwise comparisons with the MASJ population were significant. Similarly, most comparisons involving the WABA population were significant, except for that with the MAFC population. Pairwise  $G'_{ST}$  values among *P. oviforme* populations ranged from 0.045 to 0.116 (average = 0.075), with two of three comparisons significant (exception CUMB/TENN). Pairwise  $G'_{ST}$  values between species showed little to moderate differentiation, ranging from 0.013 (MAFC/TENN) to 0.101, HOLS/SCIO; average = 0.053). Pairwise estimates of allelic differentiation (Jost's *D*) showed greater genetic differentiation than fixation measures ( $G'_{ST}$ ), indicating that the most common alleles were often not shared between populations. The two measures of differentiation were generally concordant, with low or negative values between populations from the mainstem ALLE, highest

TABLE 4 Population pairwise differentiation (fixation index  $G'_{st}$ , below diagonal) and Jost's  $D$  (allelic differentiation, above diagonal) for *Pleurobema clava* and *Pleurobema oviforme*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
MASJ	MAFC	ALHS	ALWH	ALWB	ALFR	ALFC	ELK	TYGT	WABA	SCIO	GRN	CUMB	TENN	HOLS	
<i>P. clava</i>															
P1	-	0.304	0.390	0.333	0.326	0.366	0.410	0.430	0.234	0.124	0.584	0.464	0.320	0.481	0.771
2	<b>0.024</b>	-	0.224	0.295	0.269	0.279	0.298	0.426	0.393	-0.015	0.513	0.403	0.461	0.250	0.679
3	<b>0.031</b>	<b>0.017</b>	-	0.033	-0.036	-0.015	0.035	0.159	0.029	0.244	0.362	0.325	0.443	0.572	0.614
4	<b>0.028</b>	<b>0.024</b>	<b>0.003</b>	-	-0.007	-0.017	0.113	0.122	0.039	0.244	0.326	0.327	0.375	0.532	0.618
5	<b>0.025</b>	<b>0.020</b>	-0.001	-0.003	-	-0.016	0.089	0.090	-0.044	0.203	0.268	0.355	0.380	0.509	0.626
6	<b>0.028</b>	<b>0.021</b>	-0.001	-0.001	-0.001	-	0.102	0.111	0.015	0.223	0.224	0.311	0.415	0.507	0.636
7	<b>0.037</b>	<b>0.027</b>	<b>0.011</b>	0.003	0.008	0.009	-	0.228	0.199	0.303	0.454	0.359	0.448	0.708	0.710
8	<b>0.039</b>	<b>0.037</b>	<b>0.011</b>	0.014	0.008	0.010	0.023	-	0.080	0.361	0.328	0.167	0.423	0.593	0.645
9	<b>0.019</b>	<b>0.028</b>	<b>0.003</b>	0.002	0.001	0.001	0.018	0.007	-	0.236	0.405	0.384	0.300	0.598	0.580
10	<b>0.010</b>	-0.001	<b>0.019</b>	0.018	0.014	0.016	0.025	0.030	0.019	-	0.404	0.361	0.296	0.456	0.693
11	<b>0.056</b>	0.045	<b>0.033</b>	0.035	0.025	0.021	0.049	0.036	0.037	0.036	-	0.404	0.361	0.296	0.456
12	<b>0.046</b>	0.035	0.034	0.032	0.034	0.030	0.039	0.019	0.036	0.033	0.041	-	0.331	0.275	0.643
13	<b>0.036</b>	0.042	<b>0.043</b>	0.048	0.041	0.045	0.053	0.049	0.030	0.030	0.040	0.040	-	0.558	0.774
14	<b>0.036</b>	0.013	<b>0.042</b>	0.042	0.037	0.037	0.058	0.047	0.037	0.030	0.029	0.022	0.045	-	0.527
15	<b>0.093</b>	<b>0.082</b>	<b>0.076</b>	<b>0.073</b>	<b>0.074</b>	<b>0.075</b>	<b>0.093</b>	<b>0.085</b>	<b>0.071</b>	<b>0.079</b>	<b>0.101</b>	<b>0.092</b>	<b>0.116</b>	<b>0.064</b>	-

Note: Bolded values represent significant pairwise differentiation after 9999 permutations ( $p \leq .05$ ).<sup>a</sup>

<sup>a</sup>Color coding of cells indicates degree of genetic differentiation after Hartl and Clark (1997): Gray indicates no genetic differentiation ( $G'_{st} = 0.00$ ); green indicates little genetic differentiation ( $G'_{st} < 0.05$ ); and yellow indicates moderate genetic differentiation ( $G'_{st} > 0.05 - 0.15$ ).



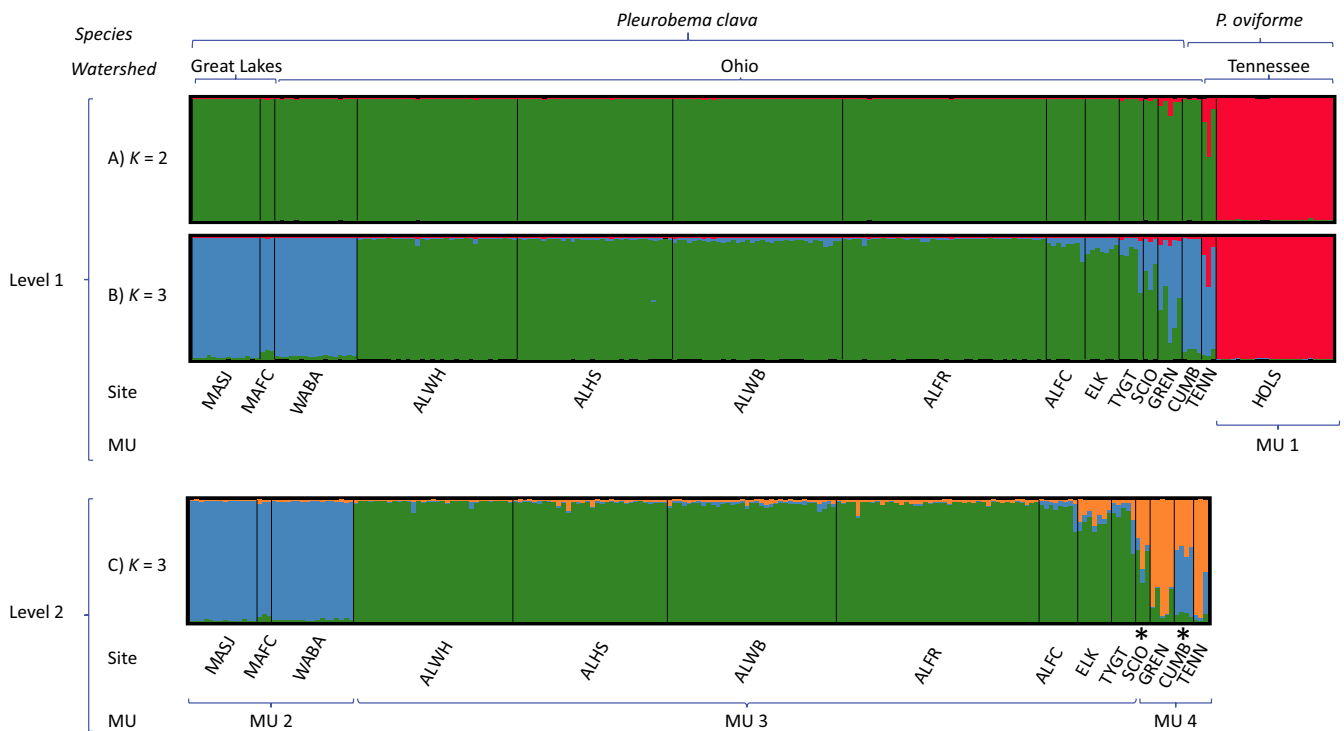
estimates in comparisons involving the HOLS population (0.456–0.774), and intermediate estimates involving MASJ and MAFC populations (0.224–0.584), with an exception of the WABA population, which were lower (0.124 and –0.015 between the MASJ and MAFC populations, respectively). Pairwise  $D$  estimates between *P. clava* populations ranged from slightly negative (ALLE, MAFC/WABA, (ALWB)/TYGT, concordant with  $G'_{ST}$  estimates) to 0.584 (MASJ/SCIO) and averaged 0.244. Estimates of  $D$  between *P. oviforme* populations ranged from 0.558 (CUMB/TENN) to 0.774 (CUMB/HOLS) and averaged 0.620. Pairwise  $D$  values between species ranged from 0.250 (MAFC/TENN) to 0.771 (MASJ/HOLS) and averaged 0.500.

Including 15 populations at the first hierarchical level of analysis, the best-supported number of clusters detected by STRUCTURE was equivocal. The presence of two clusters ( $K = 2$ ) was suggested as the most likely number of populations by the  $\Delta K$  method (Evanno et al., 2005), in which *P. oviforme* from the HOLS formed a unique genetic cluster while the remaining *P. oviforme* and *P. clava* samples formed a second cluster (Figure 7 Level 1; Figure S1(A)). At  $K = 3$ , suggested as the most likely number of populations by the mean likelihood of the probability (Mean  $\text{LnP}(X|K)$ ) (Figure 7 Level 1, Figure S1(B)), the HOLS MLGs formed one cluster, the ALLE/ELK/TYGT formed a second cluster, and the MASJ/MAFC/WABA/CUMB/TENN formed a third cluster with the SCIO and GREN

populations appearing genetically similar to both the second and third clusters.

The second hierarchical analysis included 14 populations, excluding the unique HOLS population (Cluster or MU1). Both methods suggested  $K = 3$  (Figure S1(C–D)), with clusters containing: MASJ, MAFC, and WABA (MU2); the ALLE, TYGT, and ELK (MU3); and GREN and TENN (MU4). Note that the SCIO population appeared admixed between MU3 and MU4, while the CUMB appeared admixed between MU2 and MU4. No additional structuring was detected in subsequent STRUCTURE runs on the three clusters detected at Level 2 (not shown).

Sub-populations with small sample size tend to merge in STRUCTURE analyses (Puechmaile, 2016), so the analysis was re-run including five populations with even sample sizes of 17 individuals each (WABA, ALWH, ALHS, ALWB, HOLS). Corroborating results from the full dataset, two genetic clusters were detected at the first hierarchical level, distinguishing the HOLS from other populations (Figure S2(A–C)), and the WABA population formed an independent cluster from the ALLE populations at a second hierarchical level (Figure S2(D–F)). F-statistics produced by AMOVA were examined to compare the strength of separation between clusters, but not to test for significance since both tests used the same dataset (Meirmans, 2015). The Level 1  $K = 2$  clustering producing the



**FIGURE 7** Results from hierarchical STRUCTURE analysis for 15 *Pleurobema* populations. Level 1 of the analysis included all 15 collections with average probability of membership graphs showing individual assignments to either  $K = 2$  (a,  $\Delta K$ , Evanno et al., 2005) or  $K = 3$  (b, Mean  $\text{LnP}(X|K)$ , Pritchard & Wen, 2004) genetic clusters. A second hierarchical analysis included 14 collections and suggested  $K = 3$  genetic clusters (c, Level 2). At each level of the analysis, collections were grouped based upon similar Q scores for the next level of analysis. In all graphs, individuals and populations (SITE, see Table 1 for population codes) are shown along the x-axis and assignment to populations along the y-axis. Species and watersheds are given at the top and suggested management units (MU) at bottom, with \* indicating intermediate Q scores, in which case the collection was included in both possible clusters. No additional structuring was detected within the 3 genetic clusters identified at Level 2

highest  $F_{ST}$  (0.077; Table S9D) confirming that the HOLS was the most divergent cluster, while  $K = 3$  was intermediate ( $F_{ST} = 0.048$ ; Table S9 E) and Level 2  $K = 3$  produced the lowest estimate of fixation ( $F_{ST} = 0.045$ ; Table S9F).

Results of applying the DAPC analysis including all 15 collections (Figure 8a) were consistent with those of STRUCTURE analysis (Figure 7), indicating the *P. oviforme* HOLS was the most differentiated population, while there was substantial overlap between sites from the ALLE, ELK, and TYGT populations. The remaining populations overlapped slightly yet were spread along the y-axis, with the SCIO, GREN, and the MAFC closest to the ALLE group. The placement of the CUMB population of *P. oviforme* suggested genetic similarities with the MASJ and WABA populations, while the TENN *P. oviforme* appeared similar to both the GREN and SCIO *P. clava* populations. A second DAPC analysis was performed on 14 populations after removing the divergent HOLS population (Figure 8b). Again, overlap between sites from the ALLE, ELK, and TYGT populations was apparent, but the CUMB, MASJ, and TENN populations appeared more differentiated. Genetic similarities were suggested between the WABA and SCIO populations, as well as the ELK/GREN/MAFC populations.

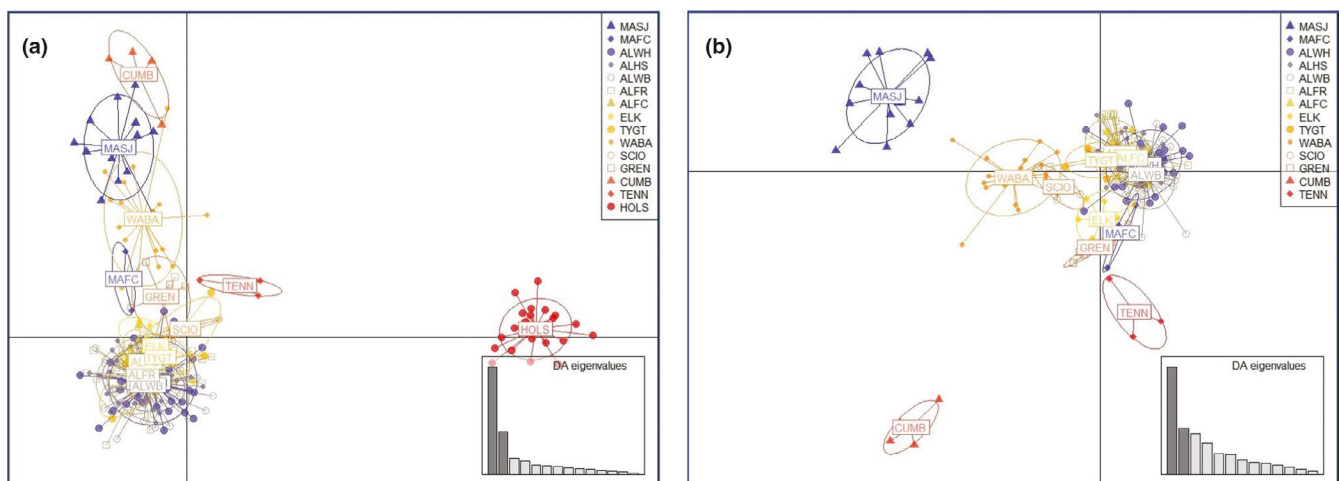
To investigate evolutionary relationships among allele frequencies, pairwise Nei's  $D_A$  distances were illustrated as a mid-point rooted neighbor-joining dendrogram (Figure S3). Most population groupings were concordant with those from the STRUCTURE analyses (Figure 7). The clade including the MASJ/MAFC/WABA populations corresponds to MU2 in the STRUCTURE graph (Figure 7c), and the clade including ALLE plus the ELK/TYGT populations corresponds to MU3 (Figure 7c). Also concordant with the STRUCTURE analysis, the SCIO and CUMB populations were intermediate between the MU groupings and basal to MU2 and MU3.

A Mantel test revealed a significant correlation between geographic river distances and genetic distance (Nei's  $D$ ,  $r = .5851$ ,

$p = .004$ , Figure S4), suggesting a stepping-stone model of dispersal (Kimura & Weiss, 1964).

## 4 | DISCUSSION

Given the long history of taxonomic uncertainty surrounding *P. clava* and *P. oviforme*, the purpose of our study was to test the taxonomic status and infer the population genetic structure of these two nominal species. Specifically, we used sequences of two mtDNA genes and morphometrics analysis of external shell measurements to test the hypotheses that these two taxa are either separate species with parapatric distributions or a single widely distributed species throughout the Ohio River, Tennessee River, and lower Great Lakes watersheds. Additionally, to help guide future conservation and management efforts, we utilized data on variation at 13 DNA microsatellite loci to characterize population genetic structure under a null hypothesis that populations represent a single readily interbreeding, undifferentiated unit with shared adaptations and a common evolutionary trajectory (Vignieri et al., 2006). While synonymy of *P. clava* and *P. oviforme* has been suspected for nearly a century based on similar shell morphologies (Ortmann, 1925), only recently have molecular data been applied to address species boundaries within the clubshell species complex, where phylogenetics and species delimitation analyses using *COI* sequence data suggested that *P. clava* and *P. oviforme* were conspecific (Campbell & Lydeard, 2012; Campbell et al., 2005, 2008; Inoue et al., 2018). Here, we have increased sampling of both putative species and coverage of geographic ranges, as well as evaluated relationships using morphometrics and multiple mtDNA genes (*COI* and *ND1*), providing the additional geographic and genetic sampling suggested by Inoue et al. (2018). A holistic approach using both molecular and morphological data sampled across



**FIGURE 8** Discriminant Analysis of Principal Components (DAPC) scatterplots based on microsatellite genotype data from *Pleurobema clava* and *Pleurobema oviforme* collections (see Table 1 for collection location and sample sizes). Panel a presents a DAPC scatterplot of 15 collections, where 100 principal components and 14 discriminant function eigenvalues were retained. Panel b presents a DAPC scatterplot of 14 collections, omitting the HOLS collection to better visualize relationships among the remaining collections, and was produced retaining 100 principal components and 13 discriminant function eigenvalues. Note that ALLE, ELK and TYGT populations are not differentiated and are superimposed in the scatterplots

the range of the taxa of interest is needed to increase understanding of often complex relationships and to better inform conservation strategies (Jones et al., 2006; Jost et al., 2018; Pante et al., 2015). Whether these taxa represent two species or a single species with multiple MUs has important implications for their future ESA listing status and conservation management. The genetic markers and the geographic coverage of the DNA sample collections utilized in this study have analytical strengths and weaknesses and underlying methodological assumptions affecting data interpretation. Hence, we discuss the strengths and weaknesses of our study results, emphasizing the caveats and key inferences from each evidence type (i.e., geography, mtDNA, nuclear DNA, morphology) and how they may affect delineation of species, MUs, and our population genetic and phylogeographic understanding of these two species.

#### 4.1 | Mitochondrial DNA evidence for a single species

Evidence for *P. clava* and *P. oviforme* being a single species was apparent in both the *COI* and *ND1* mitochondrial datasets. First, phylogenetic analysis of each mtDNA gene revealed monophyletic relationships with minimal differentiation among intermingled haplotypes (Figures 4 and 5), providing evidence for a single species based upon the PSC. Additionally, haplotype sharing between clubshell taxa was detected at both mtDNA genes, blurring any obvious genetic lines that may be drawn between the putative species. Several common haplotypes were widespread geographically, crossing nominal species and watershed boundaries. Generally, levels of mtDNA sequence divergence observed between clubshell populations were less than interspecific divergence observed in previous studies of unionid mussels (Boyer et al., 2011; Burdick & White, 2007; Inoue et al., 2014; Johnson et al., 2018; King et al., 1999; Pfeiffer et al., 2016; Pieri et al., 2018; Roe & Lydeard, 1998; Smith et al., 2019). Sequence divergence was less than 2% in all pairwise population comparisons at mtDNA genes and was moderate between putative species (0.99 and 0.66% for *COI* and *ND1*, respectively; Tables S3 and S4), with shared haplotypes between *P. clava* and *P. oviforme* at both genes, resulting in no clear “barcode gap” (Meyer & Paulay, 2005) that would be expected between intra- and interspecific divergence comparisons. There were also cases where the intraspecific genetic variation exceeded interspecific divergence. Examples include higher divergence between the Maumee and other *P. clava* populations compared with the CUMB, HOLS, and PROC *P. oviforme* populations (*COI*; Table S1) and versus the DUCK and SCHK *P. oviforme* populations (*ND1*; Table S2).

A caveat regarding our conclusion of a single clubshell species is the imbalance in population sampling among datasets, with our *COI* dataset favoring collections of *P. clava* from the Ohio River basin, while more balanced sampling throughout the ranges of both putative species was achieved at *ND1*, especially from the HOLS and CLIN drainages. Given the broader sampling of populations in the CUMB, HOLS, and TENN drainages at the *ND1* locus, this dataset provides the most

compelling evidence for a single clubshell species; however, more balanced sampling at all genetic markers would strengthen results. It is noteworthy that an undescribed taxon from the Little River clustered separately in our phylogenetic analysis of *ND1* data (Figure 5; Schilling, 2015). This result highlights the utility of this gene region in delineating cryptic taxa when they are present in a dataset, and shows potential for undocumented cryptic diversity within the clubshell species complex requiring additional taxonomic investigation.

While numerous studies of freshwater mussels have utilized mtDNA to assess relationships at various evolutionary depths, this locus reflects only one component of evolutionary history. Low information content and potential non-neutrality (i.e., selective sweeps and introgression) bring into question sole reliance on the mtDNA locus to define evolutionary lineages (Ballard & Whitlock, 2004; Edwards & Beerli, 2000; Galtier et al., 2009; Moritz, 1994; Rubinoff & Holland, 2005). Discordance between mtDNA gene trees and species trees may be expected in recently diverged lineages, as ancestral polymorphisms may be retained, a process that may be more pronounced in structured populations (Arbogast et al., 2002; Avise, 2000; Maddison, 1997; Wakeley, 2000, 2001). The significant neutrality tests at *ND1* may be indicative of selective sweeps and retention of ancestral polymorphisms, although both selection and historical demography (e.g., population bottlenecks) must be taken into consideration (Nielsen, 2005). For clubshells, bottlenecks were not detected at microsatellite loci (Table S5), making selective sweeps the more likely scenario, therefore drawing into question the reliance on these results to discern species.

By increasing the number of loci and through joint analysis of gene trees, greater success in delimiting closely related lineages has been achieved in other studies (Dupuis et al., 2012; Fujita et al., 2012; Herrera & Shank, 2016). Additionally, well-developed mtDNA sequence divergence is not always evident even among distinguishable species. For example, Jones et al. (2006) showed that several species of *Epioblasma* occurring in the Tennessee and Cumberland River basins are diverged morphologically but are minimally diverged (<1%) at several mtDNA genes and even share some of the same haplotypes among the respective populations. There are numerous examples in the scientific literature of closely related species that are distinct morphologically, behaviorally, and ecologically, but are minimally diverged at mtDNA and even share haplotypes, including some mollusks, freshwater and marine fishes, birds, and other taxa (Avise, 2000; Echelle & Dowling, 1992; McCartney et al., 2013; Pedraza-Marrón et al., 2019). Clearly, the addition of morphological characters and nuclear loci strengthens our findings.

#### 4.2 | Morphological differences between the shell forms

Quantitatively assessing the morphological variation between *P. clava* and *P. oviforme* is challenging due to the overall similarity of their shell shape, the inherent variation in the shell morphology throughout the ranges of both species, and the potential taxonomic



confusion that has occurred in the parapatric regions of each species' distribution. The taxonomic challenges are exacerbated by the over-description of morphological variation in this species complex, which has resulted in at least 16 and 23 described species that are now considered synonyms of *P. clava* and *P. oviforme*, respectively (Graf & Cummings, 2021; Watters et al., 2009; Williams et al., 2008). Additionally, the type specimen of *Unio oviformis* is apparently lost (Graf & Cummings, 2021; Williams et al., 2008) and the type specimen of *Unio clava* is a relatively large specimen (75 mm total length) with severe growth scars on the shell disk, which is not considered a diagnostic character for the species (Lamarck, 1819; Watters et al., 2009; Williams et al., 2008).

In the early 1900s, A. E. Ortmann was very aware of the morphological similarity between *P. clava* and *P. oviforme*, even stating that they could be the same species (i.e., synonymous), but he was also aware of the considerable variation observed in the headwater versus big-river forms of *P. oviforme* in the Tennessee River system (Ortmann, 1918, 1920, 1924, 1925). He recognized the inflated big-river form as *P. oviforme* "*holstonense*," which he observed primarily in the lower and middle Tennessee River and the lower reaches of its tributaries, including the Duck and Paint Rock rivers, and he recognized the compressed headwater form as *P. oviforme* "*argenteum*," which he observed mostly in the upper Tennessee River and its tributaries. But he also collected the compressed "*argenteum*" form from the upper reaches of the Duck, Elk, and Paint Rock rivers and other smaller streams, indicating that this form also was known from the lower and middle regions of the Tennessee River system, but seemingly restricted to the headwater areas of the tributaries (Ortmann, 1920, 1924). Further, the inflated "*holstonense*" form was known from the upper Tennessee River basin, but mainly from the mainstem of the Tennessee River and the very lower reaches of the Clinch, Holston, French Broad, and Nolichucky rivers (Ortmann, 1918). Hence, some level of clinal variation in the shell morphology of *P. oviforme* from headwater to larger-river environments throughout the Tennessee River basin is well established in the historical literature and in museum records for the species. While this clinal shell variation likely occurs in the Cumberland River system as well, *P. oviforme* is primarily known from headwater areas where shell inflation differences are less pronounced. Regardless, the overall shape of what has been nominally identified as *P. oviforme* from the Cumberland River system is quite similar to that of *P. clava*. In fact, individuals collected from the lower to middle Cumberland River system occasionally have been identified as *P. clava*, contributing to confusion about each species' distribution in this river system. The observed morphological variation is likely a product of ecophenotypic plasticity, driven by the diverse geological conditions of the respective watersheds, with the predominately dolomite-limestone based Valley and Ridge province of the Powell, Clinch, and Holston valleys, the sandstone-shale strata of the Cumberland Plateau, and the igneous-metamorphic strata of the Blue-Ridge mountains drained by the Nolichucky, French Broad, Little Tennessee, and Hiwassee river valleys. In the upper Tennessee River basin, specimens identified as *P. oviforme* generally have a more rounded and

flattened shape, especially in the headwaters of the major tributaries to this region. Thus, excluding the parapatric regions of each species' distribution in the lower and middle Tennessee River and Cumberland River, the shell forms exhibited by most populations in the upper Tennessee River basin are rare in the Ohio River and Great Lakes region that is generally understood to be the distribution of *P. clava*; however, the lack of a type specimen of *U. oviformis* creates difficulty assigning a species name to the shell forms commonly observed in the Tennessee and Cumberland river systems.

We measured commonly employed morphological measurements (shell length, height, width, and weight) and the respective ratios as response variables to delineate these taxa over a broad geographic range. While results of the PERMANOVA suggested significant differences in shell shape, broad overlap in shell shapes across species was apparent in the PCA and only 3% of the variance in shell shape was explained, implying a high degree of morphological overlap between the two nominal taxa (Figure 6). However, for taxa that are morphologically very similar, a limited set of measurements may not be sufficient to discriminate them. For example, it is likely that PCA analyses utilizing these same variables to delineate numerous "pigtoe" taxa in the genera *Fusconaia*, *Pleurobema*, and *Pleuronaia* would show only minor differences and therefore a high degree of morphological overlap in shell shape. Examples of highly similar shell morphologies for species in the Tennessee River system include the following: *Fusconaia cor* and *F. cuneolus*; *Pleurobema cordatum* and *Pleurobema plenum*; and *Pleuronaia barnesiana* and *Pleurobema oviforme*. All of these taxa are well diverged genetically and have other morphological traits used to distinguish among them (Hyde et al., 2020; Schilling, 2015); therefore, similarities in shell phenotype among these taxa are likely due in part to phenotypic plasticity.

Using LDA and Random Forest analyses, our overall classification rates between the two species were 65.5% and 83.2%, respectively (Table 3). Some of the misclassifications may in part be due to the morphological similarity (e.g., width-to-length ratio) of *P. oviforme* from the Duck and Paint Rock river samples to those of *P. clava*. Additionally, we do not report herein morphological differences or classification rates among the various populations that we examined. For example, the Random Forest analysis classification rate for individuals from the North Fork Holston River is >90%. Beyett et al. (2020) showed that traditional measurements and ratio-based methods were not as precise in assigning species to genetic groups or species. Conversely, Inoue et al. (2014) showed that shell width may be very important in differentiating among species and that 2D shape (using geometric approaches) was less important. A combination of methods that use multiple landmark measurements spread uniformly across the shell and accounting for shell width may prove useful to further characterize differences between populations in the clubshell complex (Hyde et al., 2020; Schilling, 2015; Willsie et al., 2020). Our morphological analyses assessed differences between the taxa using samples collected throughout the broad geographic range of each species, and our morphometric analysis is congruent with Ortmann's (1925) assessment indicating that the morphologies of *P. clava* and *P. oviforme* may represent clinal variation of a single species.

### 4.3 | Nuclear DNA population divergence and structure

Many previous studies of unionids have detected low levels of mtDNA divergence, with minimal divergence among haplotypes sampled across geographic ranges, while microsatellites provided increased resolution of population structuring (Chong et al., 2016; Jones et al., 2006; Kelly & Rhymer, 2005; Scott et al., 2020; Zanatta & Murphy, 2006, 2007). Similarly, analyses of nuclear microsatellite markers identified low to very high levels of among-population structuring among clubshell populations. The STRUCTURE analysis (Figure 7) suggested two to four unique genetic clusters that may be considered MUs following the criteria of Moritz (1994) given the significant differences in frequencies of nuclear alleles. The most genetically unique cluster included the HOLS population in Virginia. Concordantly, the largest transition in the distribution of genetic variation in the microsatellite dataset separated the HOLS population from others (Table 4). The HOLS population had moderate pairwise  $G'_{st}$  values (0.071–0.116), but very high pairwise  $D_{est}$  estimates (0.456–0.774; Table 4), formed a unique cluster at all tested  $K$  values (Figure 7), and clearly separated from other populations in the DAPC analysis (Figure 8a). These divergence values exceeded known values observed between other *Pleurobema* species; for example, Jones, Johnson, et al. (2015), utilizing similar DNA microsatellite loci, reported  $D$  values of 0.24 between sympatric populations of *P. cordatum* and *P. plenum* in the Green River, KY, and 0.42 between allopatric populations of *P. cordatum* in the Green River, KY, and *P. plenum* in the Clinch River, TN. A second genetic cluster included the ALLE, TYGT, and ELK clubshell populations of the upper Ohio Basin. A close genetic relationship among these populations also was supported in the DAPC (Figure 8) and neighbor-joining (Figure S3) analyses. The third genetic cluster included the two Maumee populations (MASJ and MAFC) as well as the Tippecanoe (WABA), which was supported by the DAPC (Figure 8) and neighbor-joining (Figure S3) analyses. A potential fourth cluster included the GREN and LTEN populations (Figure 7c), although the genetic similarity of these populations is suggestive but less clear in the DAPC (Figure 8) and neighbor-joining (Figure S3) analyses, and sample sizes were low. Therefore, four genetic clusters or MUs (Figure 7) are the best-supported and biologically relevant grouping based on our limited geographic sampling of *P. oviforme*, where the SCIO and GREN populations showed admixture with the second and third clusters, while the CUMB and TENN showed admixture with the third (Maumee and WABA) cluster.

A significant IBD pattern (Figure S4) indicates a strong correlation between genetic divergence and increasing geographic distance, supporting a stepping-stone model of gene flow (Kimura & Weiss, 1964) rather than dispersed and divergent. Isolation-by-distance patterns have been supported within several unionid species in North America (Berg et al., 1998, 2007; Elderkin et al., 2008; Ferguson et al., 2013; Inoue et al., 2013; Kelly & Rhymer, 2005; Rowe & Zanatta, 2015) and Europe (Zieritz et al., 2010); thus, the stepping-stone model of gene flow may often be appropriate for

stream-dwelling organisms with clumped distributions (Elderkin et al., 2008). The microsatellite markers also demonstrated finer-scale divergence, such as that between populations in the mainstem ALLE River and the French Creek tributary (Table 4).

Significant caveats regarding our analysis of population structuring should be recognized. First, increased sample representation of known populations within the CUMB and HOLS drainages as well as the Clinch River population, which was not analyzed using microsatellites in this study, would likely refine our knowledge of barriers to gene flow and allow greater precision in defining the geographic extent of MUs. Small sample sizes from some populations coupled with highly polymorphic loci could lead to biased estimates of genetic distances (Ruzzante, 1998) and genetic clustering (Puechmaile, 2016). Similarly, while the nuclear microsatellites adequately clustered populations, refinement of these clusters and their geographic boundaries (and suggested MUs) may be feasible through increased investigation of the nuclear genome utilizing high-throughput sequencing approaches such as Restriction Site Associated DNA sequencing (RADseq; Baird et al., 2008; Garrison et al., 2021). Considering that underlying genetic structuring may be linked to adaptive differentiation, utilizing transcriptomic single nucleotide polymorphism (SNP) data may be informative as well.

While it would be ideal to have a balanced set of samples for all datasets, the combination of molecular markers discussed here provides historical insight into the evolution of these lineages. Our analysis is strengthened by this multi-locus approach with representation from both the nuclear and mitochondrial genomes. In the future, increasing sampling and genotyping of the CUMB and HOLS populations and including assessment of the robust Clinch River population and other populations would be desirable to inform management actions. At least a dozen populations of *P. oviforme* in the Tennessee River basin have yet to be sampled and analyzed at nuclear DNA loci, which would greatly advance our understanding of genetic diversity and levels of population differentiation within and between the two species.

### 4.4 | Geologic history and phylogeographic structure

Advance and retreat of glaciers during and after the Pleistocene had dramatic effects on topography, stream characteristics, and temperature regimes that likely influenced genetic diversity and structuring among North American freshwater populations, including fishes (Bernatchez & Wilson, 1998; Morrison et al., 2006; Near et al., 2001; Strange & Burr, 1997) and unionids (Elderkin et al., 2007, 2008; Hewitt et al., 2019; Inoue et al., 2013, 2015; Jones et al., 2015; Pieri et al., 2018; Scott et al., 2020; Zanatta & Harris, 2013). Although several studies of unionids have found that structuring generally follows major hydrogeologic basins (Galbraith et al., 2015; Mock et al., 2010), our STRUCTURE results for *P. clava* and *P. oviforme* do not align fully with present-day drainages. Although several geological events could explain genetic connections that

do not align with present drainages (e.g., tectonic activity, general erosional stream capture events), one possible scenario involves ancient drainage connections that occurred as ice sheets covering the northern portion of the range retracted approximately 15,000–8,000 years ago (Dyke & Prest, 1987). For example, the close genetic relationship between the Maumee and Wabash drainages seen in shared mitochondrial haplotypes and an MLG-based genetic cluster (MU2; Figure 7) may be the result of shared connections over the past 13,600 years. Historically, the St. Joseph River drained into the Wabash (Bleuer & Moore, 1972) and was part of an “Interior Highlands” group of drainages (Mayden, 1988). More recently, this connection was cut off and the St. Joseph River was captured by the Maumee drainage (Graf, 2002; Pielou, 1991), creating a potential route for colonization to the Great Lakes watershed. Several fishes, including a known host for *P. clava* (Northern Hogsucker, *Hypentelium nigricans*, Berendzen et al., 2003), and unionids (Threeridge *Amblema plicata*, Elderkin et al., 2007; Flutedshell *Lasmigona costata*, Hewitt et al., 2016; Plain Pocketbook *Lampsilis cardium*, Hewitt et al., 2019) also show genetic evidence for a colonization pathway into the Great Lakes region via the Wabash–Maumee outlet, the western portion of which is now the Wabash drainage. Genetic similarities between other drainages also reflect ancient connections. For example, MU3 (including the ALLE, TYGT, and ELK populations; Figure 7) were included in the ancient “Upper Ohio,” while MU4 (GREN and LTEN populations) comprised populations from the “Lower Ohio” (Mayden, 1988). The two populations that appear admixed (SCIO and CUMB) lie in the center of the geographic distribution and the detected admixture may be a result of former versus present-day stream connections.

Several studies of connectivity and colonization history among unionid populations in eastern North America have suggested multiple refugia from glaciation (Beaver et al., 2019; Hewitt et al., 2019; Inoue et al., 2013). Two glacial refugia are suggested from the data presented here, including the ancient Cumberland and Tennessee Rivers. The Cumberland shows genetic connections with the Maumee and Wabash drainages (Figures 7, 8, and S3), while the Tennessee River was basal to all populations (Figure S3). Despite low divergence among populations, the *ND1* dataset provided evidence regarding colonization history. Results of a test of neutrality of the *ND1* dataset suggested a rapid expansion from a once-small effective population size. The combination of low nucleotide diversity, high haplotype diversity (since there are many haplotypes with few mutations between them), and a star-shaped haplotype network (Figure 5) suggests the signature of recent demographic expansion consistent with expansion from glacial refugia. Further, more than 90% of the *ND1* haplotypes occurred only in the UTRB, suggesting that this geographic region harbors a high amount of unique genetic variation, suggesting long-term persistence and accumulation of genetic diversity. Maintaining this adaptive capacity will be critical for any conservation efforts targeting *P. oviforme* because the Clinch River in the UTRB contains one of the few remaining demographically robust populations for this species (Fitzgerald et al., 2021).

## 5 | MANAGEMENT IMPLICATIONS AND CONCLUSIONS

Our mtDNA, nuclear DNA, and morphological datasets represent the largest to date aimed at assessing the relatedness among *P. clava* and *P. oviforme* populations. Although some discrepancies exist among the datasets, each enriches our interpretation of the complex evolutionary history underlying present-day patterns of similarities and divergence among populations (Bowen et al., 2005; Butlin et al., 2008).

Three significant findings are evident from our analyses. First, strong mtDNA divergence that is typical among well-diverged mussel species was lacking, suggesting that *P. clava* and *P. oviforme* represent a single species. Given the genetic similarity and sharing of mtDNA haplotypes across populations and the observed admixture of individuals at nuclear DNA microsatellites in key geographically intermediate populations, we consider the presence of a single species the most likely biological scenario. Second, our morphometric analyses showed overlap in shell shape and lacked congruence with current models of two species. Instead, morphological differences were more pronounced among certain population-level comparisons relative to species-level comparisons, suggesting that phenotypic differences in shell shape may be the result of ecophenotypic plasticity and not genetic differentiation. The shells, for example, from both the HOLS and CLIN populations are essentially indistinguishable morphologically but are quite distinguishable from shells of *P. clava* in the Ohio River and Great Lakes watersheds because they tend to be more compressed and rounded with umbos heavily eroded and weakly elevated with respect to the hinge. This morphology is distinctive from the original description of *U. oviformis*, where the umbos of the species were described as “prominent, not decorated.” Third, analyses of the DNA microsatellite data suggest that significant structuring exists across the broad geographic ranges of *P. clava* and *P. oviforme*, providing critical information for defining representative units to inform ESA decision making and recovery efforts. Four genetic clusters, or MUs, were detected through STRUCTURE analysis, only some of which follow present-day drainages. The most distinctive genetic cluster was comprised of the HOLS population, which ideally should be investigated further by increasing sampling and microsatellite genotyping of geographically proximate populations, such as those from the Clinch River and its tributaries, which likely comprise the largest population of the species range-wide. If CLIN mussels are closely related to HOLS, these two populations and perhaps others in the UTRB (e.g., French Broad (FRBR), Hiwassee (HIWA), Little River, Nolichucky (NOLI), PROC, SCHK) would form at a minimum a large MU. Alternatively, sampling intermediate populations may erode genetic clusters through the inclusion of genetically admixed individuals, which has been shown in groups characterized by strong patterns of isolation by distance (Mason et al., 2020).

Our data will aid management efforts to assess how the synonymy of *P. clava* and *P. oviforme* may impact extinction risk assessments and can be used to inform strategies for translocating individuals

and augmenting populations. Because *P. oviforme* is currently being considered for ESA protection and *P. clava* is listed as federally endangered, our findings suggesting that these are the same species have direct implications for upcoming ESA decisions. Synonymizing these species may remove the need to determine an individual status for *P. oviforme* and require that the current status and recovery strategies for *P. clava* be re-evaluated based on the combined distributions. However, if future genetic, morphological, life history, and breeding studies demonstrate that *P. clava* and *P. oviforme* or some of the geographically isolated populations are different species from each other, then synonymizing these two taxa now could have major negative consequences for their protection and management. Based upon genetic structuring and diversity of multiple unionid species, Galbraith et al. (2015) proposed translocating and augmenting populations within watersheds as a prudent management strategy. However, the four MUs detected in our analyses do not group strictly by watershed, but more closely align with ancient (Pleistocene) river connections. Hence, until additional genetic and morphological data are available to inform species boundaries between *P. clava* and *P. oviforme*, and reproductive data are available to inform breeding outcomes and the fitness of progeny produced from broodstock out-crossings, we suggest that population management could occur and be restricted to the MU level. Finally, high levels of genetic diversity were observed using both nuclear microsatellites as well as mtDNA haplotypes in most populations. Since both abundance and genetic diversity are important when considering the status of imperiled populations, high diversity is a positive sign for long-term stability and potential to adapt to changing environments. Similarly, little to no evidence of population bottlenecks suggests that population sizes were stable historically. Therefore, there may not be an immediate need to augment populations solely for the purpose of enhancing genetic diversity. Along with the increased sampling at nuclear loci for *P. oviforme* populations suggested above to further define the boundaries among evolutionary lineages, host-fish compatibility trials and additional research regarding ecological and demographic characteristics may guide decisions regarding taxonomy.

## ACKNOWLEDGMENTS

We give special thanks to Daniel Schilling for his field efforts and contributions to our dataset. We also wish to thank the following colleagues for sample donation: Kody Kuehnl, Jacob Culp, Wendell Haag, Steve Fraley, Art Bogan, Brant Fisher, Janet Clayton, Barbara Douglas, Rita Villella Bumgardner, Glenn Nelson, and Tony Brady. Leroy Koch, Bob Butler, and Andrew Henderson (USFWS) have provided funding to support this project. Analysis of shell morphologies was enhanced by a loan of material from the McClung Museum of Natural History and Culture enabled by Gerry Dinkins and Adriane Bernadine Tafoya and Florida Museum of Natural History facilitated by John Slapcinsky and Jim Williams. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The views expressed in this article are the authors' and do not necessarily represent those of the

U.S. Fish and Wildlife Service. The participation of EMH was supported in part by the Virginia Agricultural Experiment Station under funding from the U.S. Department of Agriculture National Institute of Food and Agriculture.

## CONFLICT OF INTEREST

None declared.

## AUTHOR CONTRIBUTIONS

**Cheryl L. Morrison:** Conceptualization (equal); Data curation (lead); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Resources (equal); Visualization (equal); Writing-original draft (lead); Writing-review & editing (lead). **Nathan A. Johnson:** Conceptualization (equal); Data curation (supporting); Formal analysis (equal); Funding acquisition (supporting); Investigation (equal); Methodology (equal); Resources (equal); Visualization (equal); Writing-original draft (supporting); Writing-review & editing (equal). **Jess W. Jones:** Conceptualization (equal); Data curation (supporting); Formal analysis (equal); Funding acquisition (supporting); Investigation (equal); Methodology (equal); Resources (equal); Visualization (equal); Writing-original draft (supporting); Writing-review & editing (supporting). **Michael S. Eackles:** Conceptualization (supporting); Data curation (supporting); Formal analysis (supporting); Investigation (equal); Methodology (supporting); Writing-original draft (supporting); Writing-review & editing (supporting). **Aaron W. Aunins:** Formal analysis (supporting); Visualization (supporting); Writing-original draft (supporting); Writing-review & editing (supporting). **Daniel B. Fitzgerald:** Formal analysis (supporting); Methodology (supporting); Visualization (supporting); Writing-original draft (supporting); Writing-review & editing (supporting). **Eric M. Hallerman:** Funding acquisition (supporting); Investigation (supporting); Methodology (supporting); Resources (supporting); Visualization (supporting); Writing-review & editing (supporting). **Tim L. King:** Conceptualization (equal); Funding acquisition (supporting); Investigation (supporting); Methodology (equal); Resources (equal).

## DATA AVAILABILITY STATEMENT

DNA sequences: GenBank accessions COI MT991776–MT991982; ND1: MW005982–MW006051. Sampling locations, DNA sequence alignments, morphological data, and microsatellite genotypes are available in Morrison et al., 2021 at USGS ScienceBase <https://doi.org/10.5066/P928BVHR>.

## ORCID

Cheryl L. Morrison  <https://orcid.org/0000-0001-9425-691X>

Nathan A. Johnson  <https://orcid.org/0000-0001-5167-1988>

Jess W. Jones  <https://orcid.org/0000-0001-5623-7945>

Michael S. Eackles  <https://orcid.org/0000-0001-5624-5769>

Aaron W. Aunins  <https://orcid.org/0000-0001-5240-1453>

Daniel B. Fitzgerald  <https://orcid.org/0000-0002-3254-7428>

Eric M. Hallerman  <https://orcid.org/0000-0002-5839-858X>



## REFERENCES

- Agrell, I. (1948). The shell morphology of some Swedish unionids as affected by ecological conditions. *Arkiv För Zoologi*, 41, 1–30.
- Arbogast, B. S., Edwards, S. V., Wakeley, J., Beerli, P., & Slowinski, J. B. (2002). Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics*, 33, 707–740.
- Avise, J. C. (1994). *Molecular markers, natural history, and evolution*. Chapman & Hall.
- Avise, J. C. (2000). *Phylogeography: The history and formation of species*. Harvard University Press.
- Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, 3(10), e3376.
- Ballard, J. W. O., & Whitlock, M. C. (2004). The incomplete natural history of mitochondria. *Molecular Ecology*, 13, 729–744.
- Beaver, C., Woolnough, D. A., & Zanatta, D. T. (2019). Assessment of genetic diversity and structure among populations of the freshwater mussel *Epioblasma triquetra* in the Laurentian Great Lakes drainage. *Freshwater Science*, 38, 527–542.
- Berendzen, P. B., Simons, A. M., & Wood, R. M. (2003). Phylogeography of the northern hogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): Genetic evidence for the existence of the ancient Teays River. *Journal of Biogeography*, 30, 1139–1152.
- Berg, D. J., Cantonwine, E. G., Hoeh, W. R., & Guttman, S. I. (1998). Genetic structure of *Quadrula quadrula* (Bivalvia: Unionidae): Little variation across large distances. *Journal of Shellfish Research*, 17, 1365–1373.
- Berg, D. J., Christian, A. D., & Guttman, S. I. (2007). Population genetic structure of three freshwater mussel (Unionidae) species within a small stream system: Significant variation at local spatial scales. *Freshwater Biology*, 52, 1427–1439. <https://doi.org/10.1111/j.1365-2427.2007.01756.x>
- Berg, D. J., Haag, W. R., Guttman, S. I., & Sickel, J. B. (1995). Mantle biopsy: A technique for nondestructive tissue-sampling of freshwater mussels. *Journal of the North American Benthological Society*, 14, 577–581.
- Bernatchez, L., & Wilson, C. C. (1998). Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, 7, 431–452. <https://doi.org/10.1046/j.1365-294x.1998.00319.x>
- Beyett, T. W., McNichols-O'Rourke, K., Morris, T. J., & Zanatta, D. T. (2020). Use of morphometric analyses and DNA barcoding to distinguish *Truncilla donaciformis* and *Truncilla truncata* (Bivalvia: Unionidae). *Freshw. Mollusk Biol. Conserv.*, 23, 99–108. <https://doi.org/10.31931/fmbc.v23i2.2020.99-108>
- Bleuer, N. K., & Moore, M. C. (1972). Glacial stratigraphy of the Fort Wayne, Indiana, area and the draining of glacial Lake Maumee. *Proceedings of the Indiana Academy of Science*, 81, 195–209.
- Boutin-Ganache, I., Raposo, M., Raymond, M., & Deschepper, C. F. (2001). M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different sizing methods. *BioTechniques*, 31, 24–28.
- Bowen, B. W., Bass, A. L., Soares, L., & Toonen, R. J. (2005). Conservation implications of complex population structure: Lessons from the loggerhead turtle (*Caretta caretta*). *Molecular Ecology*, 14, 2389–2402. <https://doi.org/10.1111/j.1365-294X.2005.02598.x>
- Boyer, S. L., Howe, A. A., Juergens, N. W., & Hove, M. C. (2011). A DNA-barcoding approach to identifying juvenile freshwater mussels (Bivalvia: Unionidae) recovered from naturally infested fishes. *Journal of the North American Benthological Society*, 30, 182–194. <https://doi.org/10.1899/10-004.1>
- Burdick, R. C., & White, M. M. (2007). Phylogeography of the Wabash pigtoe, *Fusconaia flava* (Rafinesque, 1820) (Bivalvia: Unionidae). *Journal of Molluscan Studies*, 73, 367–375.
- Burlakova, L. E., Campbell, D. C., Karatayev, A. Y., & Barclay, D. (2012). Distribution, genetic analysis and conservation priorities for rare Texas freshwater molluscs in the genera *Fusconaia* and *Pleurobema* (Bivalvia: Unionidae). *Aquat. Biosystems*, 8, 12. <https://doi.org/10.1186/2046-9063-8-12>
- Butlin, R. K., Galindo, J., & Grahame, J. W. (2008). Sympatric, parapatric or allopatric: The most important way to classify speciation? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363, 2997–3007.
- Campbell, D. C., Johnson, P. D., Williams, J. D., Rindsberg, A. K., Serb, J. M., Small, K. K., & Lydeard, C. (2008). Identification of 'extinct' freshwater mussel species using DNA barcoding. *Molecular Ecology Resources*, 8, 711–724.
- Campbell, D. C., & Lydeard, C. (2012). Molecular systematics of *Fusconaia* (Bivalvia: Unionidae: Ambleminae). *American Malacological Bulletin*, 30, 1–17.
- Campbell, D. C., Serb, J. M., Buhay, J. E., Roe, K. J., Minton, R. L., & Lydeard, C. (2005). Phylogeny of North American amblemines (Bivalvia, Unionoida): Prodigious polyphyly proves pervasive across genera. *Invertebrate Biology*, 124, 131–164.
- Chong, J. P., Harris, J. L., & Roe, K. J. (2016). Incongruence between mtDNA and nuclear data in the freshwater mussel genus *Cyprogenia* (Bivalvia: Unionidae) and its impact on species delineation. *Ecology and Evolution*, 6, 2439–2452.
- Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, 9, 1657–1659.
- Cornuet, J. M., & Luikart, G. (1996). Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144, 2001–2014.
- Cummings, K. S., & Berlocher, J. M. K. (1990). The naiades or freshwater mussels (Bivalvia: Unionidae) of the Tippecanoe River, Indiana. *Malacological Review*, 23, 83–98.
- Cummings, K. S., Meyer, C. A., & Page, L. M. (1992). *Survey of the Freshwater Mussels (Mollusca: Unionidae) of the Wabash River Drainage* (pp. 201). Final Report. Illinois Natural History Survey, Center for Biodiversity. Technical Report.
- de Lamarck, J. B. P. A. (1819). Les nayades. In *Histoire naturelle des animaux sans vertèbres* (1st ed., Vol. 5, pp. 67–100). National Museum of Natural History.
- De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, 56, 879–886.
- Dean, G. W. (1890). Distribution of Unionidae in three rivers, Mahoning, Cuyahoga and Tuscarawas. *Nautilus*, 4, 20–22.
- Dirienzo, A., Peterson, A. C., Garza, J. C., Valdes, A. M., Slatkin, M., & Freimer, N. B. (1994). Mutational processes of simple sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 3166–3170. <https://doi.org/10.1073/pnas.91.8.3166>
- Dupuis, J. R., Roe, A. D., & Sperling, F. A. H. (2012). Multi-locus species delimitation in closely related animals and fungi: One marker is not enough. *Molecular Ecology*, 21, 4422–4436.
- Dussex, N., Rawlence, N. J., & Robertson, B. C. (2015). Ancient and contemporary DNA reveal a pre-human decline but no population bottleneck associated with recent human persecution in the Kea (*Nestor notabilis*). *PLoS One*, 10(2), e0118522. <https://doi.org/10.1371/journal.pone.0118522>
- Dyke, A. S., & Prest, V. K. (1987). Late Wisconsinan and Holocene history of the Laurentide Ice Sheet. *Géographie Physique Et Quaternaire*, XLI, 2, 237–263.
- Echelle, A. A., & Dowling, T. E. (1992). Mitochondrial DNA variation and evolution of the Death Valley pupfishes (Cyprinodon, Cyprinodontidae). *Evolution*, 46, 193–206.
- Edwards, S. V., & Beerli, P. (2000). Perspective: Gene divergence, population divergence, and the variance in coalescence time in

- phylogeographic studies. *Evolution*, 54, 1839–1854. <https://doi.org/10.1111/j.0014-3820.2000.tb01231.x>
- Elderkin, C. L., Christian, A. D., Metcalfe-Smith, J. L., & Berg, D. J. (2008). Population genetics and phylogeography of freshwater mussels in North America. *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia: Unionidae). *Molecular Ecology*, 17, 2149–2163.
- Elderkin, C. L., Christian, A. D., Vaughn, C. C., Metcalfe-Smith, J. L., & Berg, D. J. (2007). Population genetics of the freshwater mussel, *Amblema plicata* (Say 1817) (Bivalvia: Unionidae): Evidence of high dispersal and post-glacial colonization. *Conservation Genetics*, 8, 355–372. <https://doi.org/10.1007/s10592-006-9175-0>
- Eldredge, N., & Cracraft, J. (1980). *Phylogenetic patterns and the evolutionary process*. Columbia University Press.
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620.
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131, 479–491.
- Ferguson, C. D., Blum, M. J., Raymer, M. L., Eackles, M. S., & Krane, D. E. (2013). Population structure, multiple paternity, and long-distance transport of spermatozoa in freshwater mussel *Lampsilis cardium* (Bivalvia: Unionidae). *Freshwater Science*, 32, 267–282.
- Fitzgerald, D. B., Henderson, A. R., Maloney, K. O., Freeman, M. C., Young, J. A., Rosenberger, A. E., Kazyak, D. C., & Smith, D. R. (2021). A Bayesian framework for assessing extinction risk based on ordinal categories of population condition and projected landscape change. *Biological Conservation*, 253, 108866. <https://doi.org/10.1016/j.biocon.2020.108866>
- Folmer, O., Black, M., Hoeh, W. R., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294–299.
- Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915–925. <https://doi.org/10.1093/genetics/147.2.915>
- Fujita, M. K., Leaché, A. D., Burbrink, F. T., McGuire, J. A., & Moritz, C. (2012). Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, 27, 480–488.
- Galbraith, H. S., Zanatta, D. T., & Wilson, C. C. (2015). Comparative analysis of riverscape genetic structure in rare, threatened and common freshwater mussels. *Conservation Genetics*, 16, 845–857.
- Galtier, N., Nabholz, B., Glemin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology*, 18, 4541–4550. <https://doi.org/10.1111/j.1365-294X.2009.04380.x>
- Garrison, N. L., Johnson, P. D., & Whelan, N. V. (2021). Conservation genomics reveals low genetic diversity and multiple parentage in the threatened freshwater mussel, *Margaritifera hembeli*. *Conservation Genetics*, 22, 217–231.
- Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, 10, 305–318.
- Graf, D. L. (1997). Northern distribution of freshwater pearly mussels (Bivalvia: Unionidae) during Wisconsin deglaciation in the southern Glacial Lake Agassiz Region: A review. *American Midland Naturalist*, 138, 37–47.
- Graf, D. L. (1998). Freshwater pearly mussels: Pigtoes and Ortmann's Law. *American Conchologist*, 26(2), 20–21.
- Graf, D. L. (2002). Historical biogeography and late glacial origin of the freshwater pearly mussel (Bivalvia: Unionidae) faunas of Lake Erie, North America. In *Occasional papers on Mollusks, The Department of Mollusks, Museum of Comparative Zoology* (Vol. 6, pp. 175–211). Harvard University.
- Graf, D. L., & Cummings, K. S. (2021). *The Freshwater Mussels (Unionoida) of the World (and other less consequential bivalves)*. MUSSEL Project Web Site. <http://www.mussel-project.net/>
- Haag, W. R. (2012). *North American freshwater mussels: Natural history, ecology, and conservation*. Cambridge University Press.
- Haag, W. R., & Cicerello, R. R. (2016). *A distributional Atlas of the freshwater mussels of Kentucky*. Scientific and Technical Series 8. Kentucky State Nature Preserves Commission.
- Haag, W. R., & Williams, J. D. (2014). Biodiversity on the brink: An assessment of conservation strategies for North American freshwater mussels. *Hydrobiologia*, 735, 45–60. <https://doi.org/10.1007/s10750-013-1524-7>
- Hartl, D. L., & Clark, A. G. (1997). *Principles of population genetics* (3rd ed.). Sinauer Associates.
- Henley, W. F., Grobler, J. P., & Neves, R. J. (2006). Non-invasive method to obtain DNA from freshwater mussels (Bivalvia: Unionidae). *Journal of Shellfish Research*, 25, 975–977.
- Herrera, S., & Shank, T. M. (2016). RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution*, 100, 70–79. <https://doi.org/10.1016/j.ympev.2016.03.010>
- Hewitt, T. L., Bergner, J. L., Woolnough, D. A., & Zanatta, D. T. (2016). Phylogeography of the freshwater mussel species *Lasmigona costata*: Testing postglacial colonization hypotheses. *Hydrobiologia*, 810, 191–206.
- Hewitt, T. L., Woolnough, D. A., & Zanatta, D. T. (2019). Population genetic analyses of *Lampsilis cardium* (Bivalvia: Unionida) reveal multiple post-glacial colonization routes into the Great Lakes drainage. *American Malacological Bulletin*, 37(1), 21–34. <https://doi.org/10.4003/006.037.0103>
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35, 518–522.
- Hoang, D. T., Vinh, L. S., Flouri, T., Stamatakis, A., von Haeseler, A., & Minh, B. Q. (2018). MPBoot: Fast phylogenetic maximum parsimony tree inference and bootstrap approximation. *BMC Evolutionary Biology*, 18, 11.
- Hubitz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9, 1322–1332.
- Hudson, R. R., Boos, D. D., & Kaplan, N. L. (1992). A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*, 9, 138–151.
- Hyde, M., Hallerman, E., Santos, R., Jones, J., Varnerin, B., da Cruz Santos Neto, G., Mansur, M. C., Moraleco, P., & Callil, C. (2020). Phylogenetic assessment of freshwater mussels *Castalia ambigua* and *C. inflata* at an ecotone in the Paraguay River basin, Brazil shows that inflated and compressed shell morphotypes are the same species. *Diversity*, 12, 481.
- Inoue, K., Hayes, D. M., Harris, J. L., Johnson, N. A., Morrison, C. L., Eackles, M. S., King, T. L., Jones, J. W., Hallerman, E. M., Christian, A. D., & Randklev, C. R. (2018). The Pleurobemini (Bivalvia: Unionida) revisited: Molecular species delineation using a mitochondrial DNA gene reveals multiple conspecifics and undescribed species. *Invertebrate Systematics*, 32, 689–702. <https://doi.org/10.1071/IS17059>
- Inoue, K., Lang, B. K., & Berg, D. J. (2015). Past climate change drives current genetic structure of an endangered freshwater mussel species. *Molecular Ecology*, 24, 1910–1926.
- Inoue, K., McQueen, A. L., Harris, J. L., & Berg, D. J. (2014). Molecular phylogenetics and morphological variation reveal recent speciation in freshwater mussels of the genera *Arcidens* and *Arkansia* (Bivalvia: Unionidae). *Biological Journal of the Linnean Society*, 112, 535–545.

- Inoue, K., Monroe, E. M., Elderkin, C. L., & Berg, D. J. (2013). Phylogeographic and population genetic analyses reveal Pleistocene isolation followed by high gene flow in a wide ranging, but endangered, freshwater mussel. *Heredity*, *112*, 282–290.
- Johnson, N. A., Smith, C. H., Pfeiffer, J. M., Randklev, C. R., Williams, J. D., & Austin, J. D. (2018). Integrative taxonomy resolves taxonomic uncertainty for freshwater mussels being considered for protection under the U.S. Endangered Species Act. *Scientific Reports*, *8*, 15892. <https://doi.org/10.1038/s41598-018-33806-z>
- Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 1–15. <https://doi.org/10.1186/1471-2156-11-94>
- Jones, J. W., Johnson, N., Grobler, P., Schilling, D., Neves, R. J., & Hallerman, E. M. (2015). Endangered rough pigtoe pearly mussel: Assessment of phylogenetic status and genetic differentiation of two disjunct populations. *Journal of Fish and Wildlife Management*, *6*(2), 338–349. <https://doi.org/10.3996/052013-JFWM-036>
- Jones, J. W., Neves, R. J., Ahlstedt, S. A., & Hallerman, E. M. (2006). A holistic approach to taxonomic evaluation of two closely related endangered freshwater mussel species, the oyster mussel *Epioblasma capsaeformis* and tan riffleshell *Epioblasma florentina walkeri* (Bivalvia: Unionidae). *Journal of Molluscan Studies*, *72*, 267–283. <https://doi.org/10.1093/mollus/eyl004>
- Jones, J. W., Neves, R. J., & Hallerman, E. M. (2015). Historical demography of freshwater mussels (Bivalvia: Unionidae): Genetic evidence for population expansion and contraction during the late Pleistocene and Holocene. *Biological Journal of the Linnean Society*, *114*, 376–397. <https://doi.org/10.1111/bij.12437>
- Jost, L. (2008).  $G_{ST}$  and its relatives do not measure differentiation. *Molecular Ecology*, *17*, 4015–4026.
- Jost, L., Archer, F., Flanagan, S., Gaggiotti, O., Hoban, S., & Latch, E. (2018). Differentiation measures for conservation genetics. *Evolutionary Applications*, *11*, 1139–1148.
- Kalinowski, S. T. (2004). Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conservation Genetics*, *5*, 539–543. <https://doi.org/10.1023/B:COGE.0000041021.91777.1a>
- Kalinowski, S. T. (2005). HP-Rare 1.0: A computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes*, *5*, 187–189.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*, 587–589.
- Kelly, M. W., & Rhymer, J. M. (2005). Population genetic structure of a rare unionid (*Lampsilis cariosa*) in a recently glaciated landscape. *Conservation Genetics*, *6*, 789–802.
- Kimura, M., & Weiss, G. H. (1964). The stepping stone model of population structure and decrease of genetic correlation with distance. *Genetics*, *49*, 561–576.
- King, T. L., Eackles, M. S., Gjetvaj, B., & Hoeh, W. R. (1999). Intraspecific phylogeography of *Lasmigona subviridis* (Bivalvia: Unionidae): Conservation implications of range discontinuity. *Molecular Ecology*, *8*, S65–S78. <https://doi.org/10.1046/j.1365-294X.1999.00784.x>
- King, T. L., Eackles, M. E., & Young, C. C. (2006). Microsatellite DNA markers for assessing phylogeographic and population structure in Preble's meadow jumping mice (*Zapus hudsonius preblei*) and cross-amplification among neighbouring taxa. *Molecular Ecology Notes*, *6*, 670–673.
- King, T. L., Lubinski, B. A., & Spidle, A. P. (2001). Microsatellite DNA variation in Atlantic sturgeon (*Acipenser oxyrinchus oxyrinchus*) and cross-species amplification in the Acipenseridae. *Conservation Genetics*, *2*, 103–119.
- Kodolova, O. P., & Logvinenko, B. M. (1973). Comparison of different populations of bivalves *Unio pictorum* and *U. tumidus* (Unionidae) by systems of myogens and shell morphology. *Zoologicheskii Zhurnal*, *52*, 988–999.
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*, 1179–1191.
- Kowalewski, M., Dyreson, E., Marcot, J., Vargas, J., Flessa, K., & Hallman, D. (1997). Phenetic discrimination of biometric simpletons: Paleobiological implications of morphospecies in the lingulide brachiopod *Glottidia*. *Paleobiology*, *23*(4), 444–469.
- Labarca, C., & Paigen, K. (1980). A simple, rapid, and sensitive DNA assay procedure. *Analytical Biochemistry*, *102*, 344–352.
- Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*, 1110–1116.
- Li, Y. L., & Liu, J. X. (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources*, *18*, 176–177.
- Lovelace, R., & Ellison, R. (2018). Stplanr: A package for transport planning. Version 0.4. *The R Journal*, *10*, 7–23. <https://doi.org/10.32614/RJ-2018-053>
- Luikart, E. A., Allendorf, F. W., Cornuet, J.-M., & Sherwin, W. B. (1998). Distortion of allele frequency distribution provides a test for recent population bottlenecks. *Genetics*, *89*, 238–247.
- Luikart, G., & Cornuet, J. M. (1998). Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conservation Biology*, *12*, 228–237.
- Lydeard, C., Cowie, R. H., Ponder, W. F., Bogan, A. E., Bouchet, P., Clark, S. A., Cummings, K. S., Frest, T. J., Gargonminy, O., Herbert, D. G., Hershler, R., Perez, K. E., Roth, B., Seddon, M., Strong, E. E., & Thompson, F. G. (2004). The global decline of nonmarine mollusks. *BioScience*, *54*, 321–330.
- Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, *46*, 523–536.
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.
- Mason, N. A., Fletcher, N. K., Gill, B. A., Funk, C. A., & Zamudio, K. R. (2020). Coalescent-based species delimitation is sensitive to geographic sampling and isolation by distance. *Systematics and Biodiversity*, *18*(3), 269–280.
- Mayden, R. L. (1988). Vicariance biogeography, parsimony, and evolution in North America freshwater fishes. *Systematic Zoology*, *37*, 329–355.
- McCartney, M. A., Burton, M. L., Lima, T. G., & Rocha, L. (2013). Mitochondrial DNA differentiation between populations of black sea bass (*Centropristis striata*) across Cape Hatteras, North Carolina (USA). *Journal of Biogeography*, *40*, 1386–1398.
- Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, *24*, 3223–3231.
- Meirmans, P. G. (2020). GENODIVE version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources*, *20*, 1126–1131.
- Meyer, C. P., & Paulay, G. (2005). DNA barcoding: Error rates and comprehensive sampling. *PLoS Biology*, *3*, e422.
- Michalakis, Y., & Excoffier, L. (1996). A generic estimation of population subdivision using distances between alleles with special interest to microsatellite loci. *Genetics*, *142*, 1061–1064.
- Mock, K. E., Box, J. C., Chong, J. P., Howard, J. K., Nez, D. A., Wolf, D., & Gardner, R. S. (2010). Genetic structuring in the freshwater mussel *Anodonta* corresponds with major hydrologic basins in the western United States. *Molecular Ecology*, *19*, 569–591.
- Moritz, C. (1994). Applications of mitochondrial-DNA analysis in conservation – A critical review. *Molecular Ecology*, *3*, 401–411.



- Morrison, C. L., Johnson, N. A., Eackles, M. E., & Jones, J. (2021). *Novel genetic resources for clubshell freshwater mussels (Pleurobema clava, P. oviforme) for enhanced conservation*. U.S. Geological Survey data release. <https://doi.org/10.5066/P928BVHR>
- Morrison, C. L., Lemarié, D. P., Wood, R. M., & King, T. L. (2006). Phylogeographic analyses suggest multiple lineages of *Crystallaria asprella* (Percidae: Etheostominae). *Conservation Genetics*, 7(1), 129–147.
- Near, T. J., Page, L. M., & Mayden, R. L. (2001). Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): An additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. *Molecular Ecology*, 10, 2235–2240. <https://doi.org/10.1046/j.1365-294X.2001.01362.x>
- Nei, M. (1972). Genetic distance between populations. *American Naturalist*, 106, 283–291. <https://doi.org/10.1086/282771>
- Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press.
- Nelson, G., & Platnick, N. I. (1981). *Systematics and biogeography: Cladistics and vicariance*. Columbia University Press.
- Neves, R. J. (1993). A state-of-the-unionids address. In K. S. Cummings, A. C. Buchanan, & L. M. Koch (Eds.), *Conservation and management of freshwater mussels* (pp. 1–10). Upper Mississippi River Conservation Committee.
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274.
- Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, 39(1), 197–218.
- Ortmann, A. E. (1918). The naiades (freshwater mussels) of the upper Tennessee drainage with notes on synonymy and distribution. *Proceedings of the American Philosophical Society*, 57, 521–626.
- Ortmann, A. E. (1920). Correlation of shape and station in fresh-water mussels (Naiades). *Proceedings of the American Philosophical Society*, 59, 268–312.
- Ortmann, A. E. (1924). Distributional features of naiads in tributaries of Lake Erie. *American Midland Naturalist*, 9, 101–117.
- Ortmann, A. E. (1925). The naiad-fauna of the Tennessee River system below Walden Gorge. *American Midland Naturalist*, 9, 321–372.
- Pante, E., Puillandre, N., Viricel, A., Arnaud-Haond, S., Aurelle, D., Castelin, M., Chenail, A., Destombe, C., Forcioli, D., Valero, M., Viard, F., & Samadi, S. (2015). Species are hypotheses: Avoid connectivity assessments based on pillars of sand. *Molecular Ecology*, 24, 525–544.
- Parmalee, P. W., & Bogan, A. E. (1998). *The freshwater mussels of Tennessee* (pp. 328). University of Tennessee Press.
- Peacock, M. M., Kirchoff, V. S., & Merideth, S. J. (2002). Identification and characterization of nine polymorphic microsatellite loci in the North American pika, *Ochotona princeps*. *Molecular Ecology Notes*, 2, 360–362.
- Peakall, R., & Smouse, P. E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6, 288–295.
- Pebesma, E. (2018). Simple features for R: Standardized support for spatial vector data. Version 0.8. *The R Journal*, 10, 439–446.
- Pedraza-Marrón, C. D. R., Silva, R., Deedes, J., Van Belleghem, S. M., Mastretta-Yanes, A., Dominguez-Dominguez, O., Rivero-Vega, R. A., Lutackas, L., Murie, D., Parkyn, D., Bullock, L. H., Foss, K., Ortiz-Zuazaga, H., Naváez-Baradica, J., Acero, A., Gomes, G., & Betancur-R, R. (2019). Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation. *Proceedings of the Royal Society B: Biological Sciences*, 286, 20182924.
- Pfeiffer, J. M., Johnson, N. A., Randklev, C. R., Howells, R. G., & Williams, J. D. (2016). Generic reclassification and species boundaries in the rediscovered freshwater mussel '*Quadrula mitchelli*' (Simpson in Dall, 1896). *Conservation Genetics*, 17, 279–292.
- Pielou, E. C. (1991). *After the ice age: The return of life to glaciated North America*. University of Chicago Press.
- Pieri, A. M., Inoue, K., Johnson, N. A., Smith, C. H., Harris, J. L., Robertson, C., & Randklev, C. R. (2018). Molecular and morphometric analyses reveal cryptic diversity within freshwater mussels (Bivalvia: Unionidae) of the western Gulf coastal drainages of the USA. *Biological Journal of the Linnean Society*, 124, 261–277. <https://doi.org/10.1093/biolinnean/bly046>
- Piry, S., Luikart, G., & Cornuet, J. M. (1999). BOTTLENECK: A program for detecting recent effective population size reductions from allele data frequencies. CBGP. See <http://www1.montpellier.inra.fr/CBGP/software/Bottleneck/pub.html>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Pritchard, J. K., & Wen, W. (2004). *Documentation for structure software: Version 2* (pp. 33). Department of Human Genetics, University of Chicago.
- Puechmaille, S. J. (2016). The program Structure does not reliably recover the correct population structure when sampling is uneven: Subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16, 608–627.
- R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Raymond, M., & Rousset, F. (1995). GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86, 248–249.
- Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution*, 43, 223–225.
- Roe, K. J., & Lydeard, C. (1998). Molecular systematics of the freshwater mussel genus *Potamilus* (Bivalvia: Unionidae). *Malacologia*, 39, 195–205.
- Rowe, M. T., & Zanatta, D. T. (2015). Investigating the genetic variation and structure of a native unionid mussel in the Laurentian Great Lakes following an invasion of dreissenid mussels. *Biological Invasions*, 17, 351–364. <https://doi.org/10.1007/s10530-014-0734-x>
- Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sanchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34, 3299–3302.
- Rubinoff, D., & Holland, B. S. (2005). Between two extremes: Mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology*, 54(6), 952–961. <https://doi.org/10.1080/10635150500234674>
- Ruzzante, D. E. (1998). A comparison of several measures of genetic distance and population structure with microsatellite data: Bias and sampling variance. *Canadian Journal of Fisheries and Aquatic Science*, 55, 1–14.
- Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4, 406–425.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: A laboratory manual* (2nd ed.). Cold Spring Harbor Press.
- Schilling, D. E. (2015). *Assessment of morphological and molecular genetic variation of freshwater mussel species belonging to the genera Fusconaia, Pleurobema, and Pleuroaia in the upper Tennessee River basin* (pp. 194). Master of Science Thesis. Virginia Tech.
- Scott, M. W., Morris, T. J., & Zanatta, D. T. (2020). Population structure, genetic diversity, and colonization history of the eastern pond-mussel, *Sagittunio nasutus*, in the Great Lakes drainage. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 30, 631–646.
- Serb, J. M., Buhay, J. E., & Lydeard, C. (2003). Molecular systematics of the North American freshwater bivalve genus *Quadrula* (Unionidae):



- Ambleminae) based on mitochondrial ND1 sequences. *Molecular Phylogenetics and Evolution*, 28, 1–11. [https://doi.org/10.1016/S1055-7903\(03\)00026-5](https://doi.org/10.1016/S1055-7903(03)00026-5)
- Shea, C. P., Peterson, J. T., Wisniewski, J. M., & Johnson, N. A. (2011). Misidentification of freshwater mussel species (Bivalvia: Unionidae): Contributing factors, management implications, and potential solutions. *Journal of the North American Benthological Society*, 30, 446–458.
- Smith, C. H., Johnson, N. A., Inoue, K., Doyle, R. D., & Randklev, C. R. (2019). Integrative taxonomy reveals a new species of freshwater mussel, *Potamilus streckeri* sp. nov. (Bivalvia: Unionidae): Implications for conservation and management. *Systematics and Biodiversity*, 17(4), 331–348.
- Strange, R. M., & Burr, B. M. (1997). Intraspecific phylogeography of North American highland fishes: A test of the Pleistocene vicariance hypothesis. *Evolution*, 51(3), 885–897.
- Strauss, R. E. (1985). Evolutionary allometry and variation in body form in the South American catfish genus *Corydoras* (Callichthyidae). *Systematic Biology*, 34(4), 381–396.
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595. <https://doi.org/10.1093/genetics/123.3.585>
- Takezaki, N., Nei, M., & Tamura, K. (2010). POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular Biology and Evolution*, 27(4), 747–752.
- Templeton, A. R., Crandall, K. A., & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, 132, 619–633.
- Turgeon, D. D., Quinn, J. F. Jr., Bogan, A. E., Coan, E. V., Hochberg, F. G. Jr., Lyons, W. G., Mikkelsen, P. M., Neves, R. J., Roper, C. F. E., Rosenberg, G., Roth, B., Scheltema, A., Thompson, F. G., Vecchione, M., & Williams, J. D. (1998). *Common and scientific names of aquatic invertebrates from the United States and Canada: Mollusks* (2nd ed.). Special Publication No. 26. American Fisheries Society.
- U.S. Fish and Wildlife Service. 1993. *Endangered and threatened wildlife and plants; Determination of Endangered status for the Northern Riffleshell Mussel (Epioblasma torulosa rangiana) and the Clubshell Mussel (Pleurobema clava)*. <http://endangered.fws.gov/r/fr93488.html>
- U.S. Fish and Wildlife Service (USFWS) (1994). *Clubshell (Pleurobema clava) and Northern Riffleshell (Epioblasma torulosa rangiana) Recovery Plan* (pp. 68). U.S. Fish and Wildlife Service (USFWS).
- van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICROCHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535–538.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer. <https://www.stats.ox.ac.uk/pub/MASS4/>. ISBN 0-387-95457-0.
- Vignieri, S. N., Hallerman, E. M., Bergstrom, B. J., Hafner, D. J., Martin, A. P., Devers, P., Grobler, P., & Hitt, N. (2006). Mistaken view of taxonomic validity undermines conservation of an evolutionarily distinct mouse: A response to Ramey et al (2005). *Animal Conservation*, 9, 237–243.
- Wakeley, J. (2000). The effects of subdivision on the genetic divergence of populations and species. *Evolution*, 54, 1092–1101.
- Wakeley, J. (2001). The coalescent in an island model of population subdivision with variation among demes. *Theoretical Population Biology*, 59, 133–144.
- Watters, G. T. (1988). *The naiad fauna of selected streams in Ohio. I. Stillwater River of Miami River. II. Stream systems of south-central Ohio from the Little Miami River to the Hocking River, excluding the Scioto River proper* (pp. 440). Final Report to the Division of Fish and Wildlife, Indiana Department of Natural Resources.
- Watters, G. T. (1994). Form and function of unionoidean shell sculpture and shape (Bivalvia). *American Malacological Bulletin*, 11, 1–20.
- Watters, G. T., Hoggarth, M. A., & Stansbery, D. H. (2009). *The freshwater mussels of Ohio*. The Ohio State University.
- Williams, G. J. (2011). *Data mining with Rattle and R: The art of excavating data for knowledge discovery, series (Use R!)*. Springer.
- Williams, J. D., Bogan, A. E., Butler, R. S., Cummings, K. S., Garner, J. T., Harris, J. L., Johnson, N. A., & Watters, G. T. (2017). A revised list of the freshwater mussels (Mollusca: Bivalvia: Unionida) of the United States and Canada. *Freshwater Mollusk Biology and Conservation*, 20, 33.
- Williams, J. D., Bogan, A. E., & Garner, J. T. (2008). *Freshwater Mussels of Alabama and the Mobile Basin in Georgia*. University of Alabama Press.
- Williams, J. D., Warren, M. L., Cummings, K. S., Harris, J. L., & Neves, R. J. (1993). Conservation status of freshwater mussels of the United States and Canada. *Fisheries*, 18, 6–22.
- Willsie, J. A., Morris, T. J., & Zanatta, D. T. (2020). Morphometric analyses distinguish Wabash Pigtoe (*Fusconaia flava*) and Round Pigtoe (*Pleurobema sintoxia*) mussels. *Diversity*, 12, 337.
- Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, 15, 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Zanatta, D. T., & Harris, A. T. (2013). Phylogeography and Genetic Variability of the Freshwater Mussels (Bivalvia: Unionidae) Ellipse, *Venustaconcha ellipsiformis* (Conrad 1836), and Bleeding Tooth, *V. pleasii* (Marsh 1891). *American Malacological Bulletin*, 31, 267–279.
- Zanatta, D. T., & Murphy, R. W. (2006). Development and characterization of microsatellite markers for the endangered northern riffleshell mussel *Epioblasma torulosa rangiana* (Bivalvia: Unionidae). *Molecular Ecology Notes*, 6, 850–852. <https://doi.org/10.1111/j.1471-8286.2006.01367.x>
- Zanatta, D. T., & Murphy, R. W. (2007). Range-wide population genetic analysis of the endangered northern riffleshell mussel, *Epioblasma torulosa rangiana* (Bivalvia: Unionida). *Conservation Genetics*, 8, 1393–1404. <https://doi.org/10.1007/s10592-007-9290-6>
- Zieritz, A., Hoffman, J. I., Amos, W., & Aldridge, D. C. (2010). Phenotypic plasticity and genetic isolation-by-distance in the freshwater mussel *Unio pictorum* (Mollusca: Unionida). *Evolutionary Ecology*, 24, 923–938. <https://doi.org/10.1007/s10682-009-9350-0>

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Morrison, C. L., Johnson, N. A., Jones, J. W., Eackles, M. S., Aunins, A. W., Fitzgerald, D. B., Hallerman, E. M., & King, T. L. (2021). Genetic and morphological characterization of the freshwater mussel clubshell species complex (*Pleurobema clava* and *Pleurobema oviforme*) to inform conservation planning. *Ecology and Evolution*, 11, 15325–15350. <https://doi.org/10.1002/ece3.8219>