Journal of Heredity, 2024, **115**, 72–85
https://doi.org/10.1093/jhered/esad075
Advance access publication 28 November 2023
Original Article

American Genetic Association

OXFORD

Original Article

Glacial vicariance and secondary contact shape demographic histories in a freshwater mussel species complex

Nathan A. Johnson[1,*], Andrew R. Henderson[2], Jess W. Jones[3], Caitlin E. Beaver[1], Steven A. Ahlstedt[4], Gerald R. Dinkins[4], Nathan L. Eckert[5], Mark J. Endries[2], Jeffrey T. Garner[6], John L. Harris[7], Paul D. Hartfield[8], Don W. Hubbs[9], Timothy W. Lane[10], Monte A. McGregor[11], Kendall R. Moles[12], Cheryl L. Morrison[13], Matthew D. Wagner[8], James D. Williams[14], Chase H. Smith[15]

[1]U.S. Geological Survey, Wetland and Aquatic Research Center, Gainesville, FL, United States
[2]U.S. Fish and Wildlife Service, Ecological Services, Asheville, NC, United States
[3]U.S. Fish and Wildlife Service, Virginia Tech University, Blacksburg, VA, United States
[4]McClung Museum of Natural History and Culture, University of Tennessee, Knoxville, TN, United States
[5]U.S. Fish and Wildlife Service, Neosho National Fish Hatchery, Neosho, MO, United States
[6]Alabama Division of Wildlife and Freshwater Fisheries, Florence, AL, United States
[7]Arkansas State University Museum of Zoology, Jonesboro, AR, United States
[8]U.S. Fish and Wildlife Service, Ecological Services, Jackson, MS, United States
[9]DJH Environmental Services, Camden, TN, United States
[10]Virginia Department of Wildlife Resources, Marion, VA, United States
[11]Kentucky Department of Fish and Wildlife Resources, Frankfort, KY, United States
[12]Arkansas Game and Fish Commission, Benton, AR, United States
[13]U.S. Geological Survey, Eastern Ecological Science Center, Kearneysville, WV, United States
[14]Florida Museum, University of Florida, Gainesville, FL, United States
[15]Department of Integrative Biology, University of Texas at Austin, Austin, TX, United States

[*]Corresponding author: U.S. Geological Survey, Wetland and Aquatic Research Center, 7920 NW 71st Street, Gainesville, FL 32653, United States.
Email: najohnson@usgs.gov

Corresponding Editor: William Murphy

Abstract

Characterizing the mechanisms influencing the distribution of genetic variation in aquatic species can be difficult due to the dynamic nature of hydrological landscapes. In North America's Central Highlands, a complex history of glacial dynamics, long-term isolation, and secondary contact have shaped genetic variation in aquatic species. Although the effects of glacial history have been demonstrated in many taxa, responses are often lineage- or species-specific and driven by organismal ecology. In this study, we reconstruct the evolutionary history of a freshwater mussel species complex using a suite of mitochondrial and nuclear loci to resolve taxonomic and demographic uncertainties. Our findings do not support *Pleurobema rubrum* as a valid species, which is proposed for listing as threatened under the U.S. Endangered Species Act. We synonymize *P. rubrum* under *Pleurobema sintoxia*—a common and widespread species found throughout the Mississippi River Basin. Further investigation of patterns of genetic variation in *P. sintoxia* identified a complex demographic history, including ancestral vicariance and secondary contact, within the Eastern Highlands. We hypothesize these patterns were shaped by ancestral vicariance driven by the formation of Lake Green and subsequent secondary contact after the last glacial maximum. Our inference aligns with demographic histories observed in other aquatic taxa in the region and mirrors patterns of genetic variation of a freshwater fish species (*Erimystax dissimilis*) confirmed to serve as a parasitic larval host for *P. sintoxia*. Our findings directly link species ecology to observed patterns of genetic variation and may have significant implications for future conservation and recovery actions of freshwater mussels.

Key words: biogeography, Central Highlands, endangered species, glacial dynamics, *Pleurobema*, Unionidae

Introduction

Understanding contemporary and historical mechanisms influencing the spatial distributions of species is a central goal of biogeography (Avise 2000). The increasing availability of DNA sequencing technology has led to modern studies routinely using genetic data to explore mechanisms

Received June 20, 2023; Accepted November 27, 2023

that have influenced species' demographic histories. In natural populations, genetic divergence is often correlated with geographic distance, conforming to the isolation-by-distance (IBD) model (Wright 1943). However, populations can deviate from the IBD model based on several processes, including vicariance, secondary contact, or infrequent, and unexpected long-distance dispersal (Slatkin 1985). This, along with contemporary dispersal opportunities being highly restricted by human-mediated degradation of suitable habitat patches and restriction of gene flow (e.g. dams), can produce complex demographic histories that deviate from the IBD model. This is certainly the case in freshwater habitats, which are among the most imperiled ecosystems globally (Strayer and Dudgeon 2010) with an often hydrologically complex evolution due to successive cycles of climatic changes during the Pliocene and Pleistocene when the repeated connection and isolation of waterbodies occurred during glaciation and sea level changes (Galloway 2008; Galloway et al. 2011).

The inland waters of North America have been sculpted by a series of climatic and geological events. North American river drainages in the Eastern, Ouachita, and Ozark highland regions (together considered the Central Highlands) are considered hotspots of aquatic endemism and harbor the highest global diversity of freshwater gastropods, crayfishes, fishes, and mussels (Ortmann 1924, 1925; Starnes and Etnier 1986; Etnier and Starnes 1993; Lydeard and Mayden 1995; Parmalee and Bogan 1998; Crandall and Buhay 2008). It is hypothesized that aquatic biodiversity in this region has been shaped by the repeated advancement and contraction of glaciers and ice sheets through the Miocene, Pliocene, and Pleistocene (e.g. Mayden 1988; Bernatchez and Wilson 1998; Berendzen et al. 2003, 2008; Ray et al. 2006). During the Pleistocene, the Central Highlands were divided into two geographically discontinuous upland areas by the deposition of glacial alluvium, till, and loess in the Mississippi River floodplain (Thornbury 1965; Mayden 1985): 1) the Eastern Highlands, situated east of the Mississippi River, and 2) the Interior Highlands, located west of the Mississippi River (Fig. 1). The Interior Highlands were subsequently split into the Ozark and Ouachita Highlands during the formation of Arkansas River floodplain (Robison 1986; Mayden 1988; Haag 2010). The successive cycles of climatic changes and dynamic nature of hydrological landscapes in these regions have created complex demographic histories of species and their populations that do not always exhibit genealogical concordance (Avise 2000). Historical geological events have influenced freshwater biodiversity through the fragmentation and genetic divergence of populations, and subsequent secondary contact of individuals from separate refugia (e.g. Berendzen et al. 2003; Simons 2004). Responses to glacial dynamics in North American aquatic taxa vary and are mostly lineage- or species-specific (Strange and Burr 1997; Kinziger et al. 2001; Near et al. 2001; Switzer and Wood 2002; Soltis et al. 2006; Elderkin et al. 2008; Inoue et al. 2014; Jones et al. 2015b). Elucidating geographic patterns of genetic variation can therefore be complicated by the impacts of Pleistocene glaciations, especially for freshwater species with narrow ranges and dispersal capabilities that are not fully understood.

Focal system

Freshwater mussels (Bivalvia: Unionida) are a highly diverse group with approximately 1,000 species recognized globally and 388 endemic to North America (Graf and Cummings 2021). Although taxonomically diverse, this group is also one of the most threatened in North America with over 65% of the species considered of conservation concern (Haag and Williams 2014). Imperilment stems from a combination of anthropogenic impacts, including habitat modification and invasive species, and biological characteristics that make mussels sensitive to ecosystem state change (Williams et al. 1993; Haag 2012)—mussels are benthic, filter feeding ectotherms, and nearly all species rely on a parasitic larval stage that requires temporary attachment to freshwater fishes, often with a high degree of host specificity (Barnhart et al. 2008). The high level of imperilment has served as a catalyst for recent research to improve the understanding of the ecology and evolution of freshwater mussels (e.g. Smith et al. 2020; Hewitt et al. 2021; Lopes-Lima et al. 2021a; Neemuchwala et al. 2023).

Molecular data have revealed inaccuracies in morphology-based taxonomy due to high levels of intraspecific variability and interspecific convergence in morphological traits (reviewed by Lopes-Lima et al. 2021b). Although recent research incorporating molecular data has played a pivotal role in describing the diversity of the group, many species-level hypotheses primarily based on morphology remain untested. Species delineation in the tribe Pleurobemini, and specifically the genus *Pleurobema*, exemplify issues related to high levels of morphological variability and uncertain systematics. Species identification in this group is exceptionally challenging due to high levels of intraspecific variation and interspecific convergence of morphological traits (Ortmann 1920; Shea et al. 2011; Inoue et al. 2018; Olivera-Hyde et al. 2023). Recent molecular studies have been integral in resolving taxonomic uncertainty of some *Pleurobema* species (Perkins et al. 2017; Inoue et al. 2018; Morrison et al. 2021; Johnson et al. 2023), but questions remain regarding the validity and distribution of multiple species within the genus.

The Pyramid Pigtoe, *Pleurobema rubrum* (Rafinesque, 1820), and the Round Pigtoe, *Pleurobema sintoxia* (Rafinesque, 1820) were both described in the same publication with limited detail regarding type localities. "Found in the Kentucky [River]" was included in the original description of *P. rubrum* and the type specimen mentioned by Vanatta (1915) was designated as the lectotype (ANSP 20237) by Johnson and Baker (1973); however, the origin of the lectotype is unknown. Similarly, the closest mention of a type locality in the original description of *P. sintoxia* is "found in the Ohio [River]" and Johnson and Baker (1973) designated the same specimen mentioned by Vanatta (1915) as the lectotype (ANSP 20208 from the Ohio River). Previous studies using mitochondrial sequence data and external shell morphology indicate the two species are potentially conspecifics without formal taxonomic revisions or recommendations (Campbell et al. 2005; Campbell and Lydeard 2012; Jones et al. 2015a; Inoue et al. 2018; Olivera-Hyde et al. 2023). The scientific community at large still considers *P. rubrum* a valid taxon (Jones et al. 2005, 2021; Williams et al. 2008, 2017; Watters et al. 2009; Haag and Cicerello 2016) pending a comprehensive taxonomic assessment that includes genetic sampling throughout the ranges of both nominal species. *Pleurobema rubrum* and *P. sintoxia* are thought to have been historically widespread in the Mississippi Basin from Wisconsin to Louisiana, including the Eastern and Interior highlands regions (Vidrine 1993; Parmalee and Bogan 1998; Williams et al. 2008; Watters et al. 2009). The

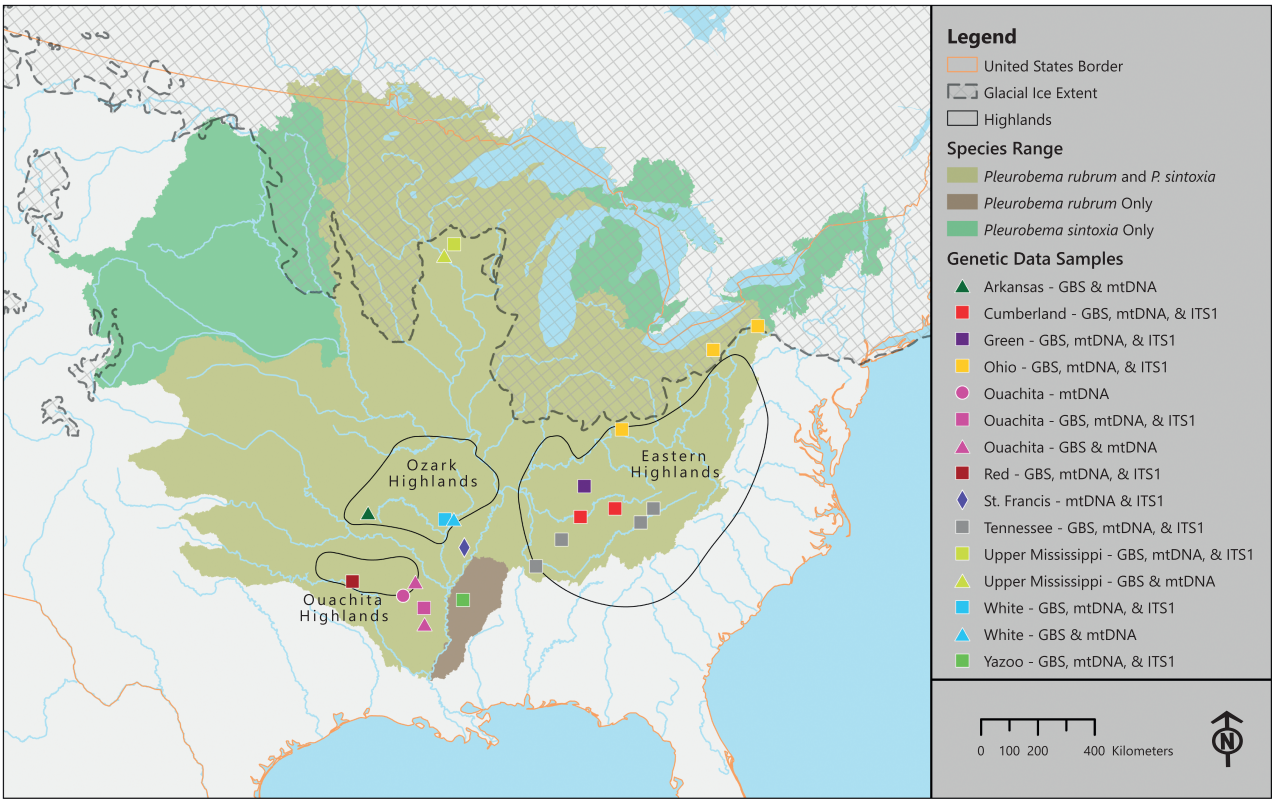

Fig. 1. Map illustrating the ranges of *P. rubrum* and *P. sintoxia* with symbol colors to indicate drainage of capture and symbol shapes to reflect the genetic data collected to represent each population. The line demarcating the extent of ice sheets during the last glacial maximum follows Ehlers et al. (2011) and approximate area of the Eastern, Ouachita, and Ozark highland regions are encircled.

two species are considered sympatric across most of their respective ranges (see Fig. 1), with the only exceptions being the eastern Great Lakes basin and upper Missouri River (only *P. sintoxia*) and the Yazoo River basin (only *P. rubrum*) (Williams et al. 2008; Watters et al. 2009; Jones et al. 2021). However, molecular and morphological similarities between the two species have led researchers to question the distribution and validity of *P. rubrum* (Ortmann 1911; Parmalee and Bogan 1998; Inoue et al. 2018; Olivera-Hyde et al. 2023). Resolving this longstanding taxonomic uncertainty became urgent when *P. rubrum* was proposed to be listed as threatened under the U.S. Endangered Species Act (USFWS 2021).

In this study, we investigate the relationships within and among extant populations of *P. rubrum* and *P. sintoxia* to resolve taxonomic uncertainties, better understand historical demographics, and help guide conservation action. Specifically, we analyzed DNA sequence data from two mitochondrial (mtDNA) genes, one nuclear DNA (nDNA) locus, and genome-wide single nucleotide polymorphisms (SNPs) generated using genotype-by-sequencing (GBS) to: 1) test the morphology-based taxonomic hypothesis that *P. rubrum* and *P. sintoxia* represent distinct species, 2) infer genetic structure across the ranges of both *P. rubrum* and *P. sintoxia*, 3) test demographic hypotheses of extant populations based on glacial dynamics, 4) synthesize possible drivers of the observed genetic variation, and 5) discuss the conservation implications of our findings. The results of our study have direct implications for conservation actions and provide another empirical example of the effects of glacial dynamics on genetic variation in aquatic taxa.

Materials and methods

Sample collection, DNA extraction, and Sanger sequencing

Our goal was to include samples representing all extant populations of *P. rubrum* and *P. sintoxia* with a focus on capturing morphological variation and broad geographic coverage across the complex mosaic of allopatric and sympatric populations (Fig. 1). By coordinating sample procurement with state agencies, regional experts, and museum curators, we were able to achieve broad geographic coverage of individuals representing the current concept of both species based on morphological identification and geographic distribution. Existing samples in museum collections were screened before additional collections were made. All collections were made with proper permissions and specimens were either non-lethally swabbed following Henley et al. (2006) or mantle clipped and released at site of capture, or vouchered in 95% non-denatured ethanol before being deposited in a public museum. Specifics on all collections are available from Johnson and Smith (2023).

DNA was extracted from DNA swabs or mantle tissue using the Qiagen PureGene extraction kit following manufacturer protocols (Qiagen; Hilden, Germany). High molecular weight DNA was ensured by visualizing isolations on a 1% agarose gel stained with Ethidium Bromide. We amplified and bidirectionally sequenced segments of two protein-coding mtDNA genes and one nDNA locus, which are commonly used in freshwater mussel species delineation studies (e.g. Campbell and Lydeard 2012; Johnson et al. 2018; Smith et al.

2018, 2019, 2021a): the mtDNA protein-coding genes *cytochrome c oxidase subunit 1* (*COX1*) and *NADH dehydrogenase subunit 1* (*ND1*), and the nDNA intron *internal transcribed spacer 1* (*ITS1*). Freshwater mussels exhibit biparental inheritance of mitochondria (Breton et al. 2011) and primers targeted the female-transmitted copies of *COX1* and *ND1*. Primers and thermal cycling conditions used for PCR and sequencing follow Johnson et al. (2018). We used Geneious v 10.0.9 (Kearse et al. 2012) to edit chromatograms and assemble consensus sequences before aligning sequences in Mesquite v 3.81 (Maddison and Maddison 2018) using MAFFT v 7.299 (Katoh and Standley 2013).

To assess genetic differentiation between the two nominal species, we generated TCS haplotype networks in PopART v1.9 (Leigh and Bryant 2015) using a concatenated alignment of *COX1* and *ND1* and a separate alignment for *ITS1*. With the DNA sequences, we first assessed species boundaries based on initial field-based identification. Next, we evaluated relationships among haplotypes based on drainage of capture to investigate the geographic distribution of genetic variation. We also calculated uncorrected pairwise genetic distances between populations using MEGA 11 (Tamura et al. 2021).

Genotype-by-sequencing

Sequencing for SNP genotyping was performed using DArTseq (DArT Pty Ltd.), which represents a combination of DArT complexity reduction methods and next generation sequencing platforms (Sansaloni et al. 2011; Kilian et al. 2012; Courtois et al. 2013; Cruz et al. 2013; Raman et al. 2014). After testing several restriction enzyme combinations for complexity reduction, the PstI–SphI combination was selected for the *P. rubrum* and *P. sintoxia* species complex. DNA samples were processed in digestion/ligation reactions following Kilian et al. (2012) but replacing a single PstI-compatible adaptor with two different adaptors corresponding to the PstI–SphI RE overhangs. The adapters were designed to include an Illumina flowcell attachment sequence, sequencing primer sequence, and barcode region (see Elshire et al. 2011).

Only PstI–SphI fragments were effectively amplified in 30 rounds of PCR using the following thermocycling conditions: 94 °C for 1 min followed by 30 cycles of 94 °C for 20 s, 58 °C for 30 s, and 72 °C for 45 s, with a final extension of 72 °C for 7 min. Equimolar amounts of amplification products from each sample of the 96-well microtiter plate were bulked and applied to c-Bot (Illumina) bridge PCR followed by sequencing on Illumina Hiseq 2500. The single read sequencing was run for 77 cycles. Sequences generated from each lane were processed using proprietary DArT analytical pipelines. In the primary pipeline the fastq files were first processed to filter out poor-quality sequences, applying more stringent selection criteria to the barcode region compared to the rest of the sequence. Filtering was performed on the raw sequences using the following parameters: Barcode region—minimum Phred pass score 30, minimum pass percentage 75; Whole read—minimum Phred pass score 10, minimum pass percentage 50.

Approximately 2,500,000 sequences per barcoded sample were identified and used in marker calling. Identical sequences were collapsed into "fastqcoll files." The fastqcoll files were "groomed" using DArT PL's proprietary algorithm which corrects low quality bases from singleton tags into correct bases using collapsed tags with multiple members as a template. The "groomed" fastqcoll files were used in the secondary pipeline for DArT PL's proprietary calling algorithms to identify SNPs and SilicoDArTs (presence/absence of restriction fragments in representation) calling algorithms (DArTsoft14). For SNP calling, all tags from all libraries included in the DArTsoft14 analysis are clustered using DArT PL's C++ algorithm at the threshold distance of 3, followed by parsing of the clusters into separate SNP loci using a range of technical parameters, including the balance of read counts for the allelic pairs. Additional selection criteria were added to the algorithm based on analysis of approximately 1,000 controlled cross populations. Testing for Mendelian distribution of alleles in these populations facilitated selection of technical parameters that effectively discriminated true allelic variants from paralogous sequences. Additionally, multiple samples were processed from DNA to allelic calls as technical replicates, and scoring consistency was used as the main selection criteria for high quality/low error rate markers. The quality of allelic calls was assured by high average read depth per locus (average across all markers was over 30 reads/locus).

Using the R package dartR v 2.7.2 (Mijangos et al. 2022), we performed additional filtering steps following similar methodologies as outlined in previous publications (Georges et al. 2018; Smith et al. 2021b). Loci with less than 100% reproducibility (see Wenzl et al. 2004) or greater than 10% missing data were removed. We then filtered individuals with greater than 10% missing data and alleles with frequencies less than 0.05 and retained the SNP with the highest degree of polymorphism at each locus to address linkage.

Genetic diversity and population structure

For all downstream GBS analyses, individuals were binned based on either nominal species or drainage of capture. We visualized the spatial distribution of genetic diversity using a principal coordinate analysis (PCoA) in dartR. To determine areas where deviations from IBD occur (e.g. Wang et al. 2010), we performed a Procrustes analysis in the R package vegan 2.5-7 (Oksanen et al. 2022). The Procrustes analysis used the first two PCs generated by PCoA, which were rotated and scaled based on sampling coordinates. We used the function "protest" to test for a significant correlation between molecular and geographical distances using 10,000 permutations and a critical value of 0.05. We further investigated the relationship between genetic and geographic distances using a Mantel test. For the Mantel test, genetic distances were averaged across all SNPs and geographic distances were calculated using a stream distance approach to better represent geographic distance among localities. Stream distances were calculated as the shortest path between sampling locations along U.S. National Hydrography flowlines (NHDPlus version 2) with reference to the Albers Equal Area North American Datum 1983 using the packages sf and stplanr (Lovelace and Ellison 2018; Pebesma 2018; R Core Team 2022). Mantel tests used 10,000 permutations and a critical value of 0.05.

We used four approaches to assess population structure in our dataset: 1) pairwise F_{ST}, 2) a discriminant analysis of principal components (DAPC), 3) the Bayesian clustering algorithm fastSTRUCTURE (Raj et al. 2014), and 4) the non-negative matrix factorization algorithm TESS3 (Caye et al. 2016). Pairwise F_{ST} was calculated between sampled populations in dartR using 1,000 bootstrap replicates. We performed DAPC in adegenet v 2.1.5 (Jombart 2008; Jombart and Ahmed 2011) on the first eight PCs and two DA eigenvalues,

which were selected based on when contributions reached a plateau. The best-fit number of clusters (K) was determined using k-means. We modeled K = 1–10 in fastSTRUCTURE, and the chooseK.py script was used to select the best K value to explain genetic structure and maximize likelihood. TESS3 incorporates geographic information into ancestry estimation, and we modeled K = 1–10 with both genotypic data and collection coordinates. Cross-validation criterion was used to select the most likely K.

Demographic history scenarios and inferences

After evaluation of genetic variation with respect to geography, we noticed the pattern of genetic variation within samples from the Eastern Highlands (i.e. Cumberland, Green, Ohio, and Tennessee River drainages) did not follow expectations of IBD. To further examine this geographic area, we modeled demographic scenarios in DIYABC-RF v 1.1.27 (Collin et al. 2021) to infer the effect of past geologic events on observed genetic variation. First, we created a population-level phylogenetic reconstruction based on GBS data using the coalescent-based approach SVDquartets (Chifman and Kubatko 2014) in PAUP* v 4.0a (Swofford 2002) with 100 bootstrap replicates for nodal support. We then created seven scenarios (Supplementary Fig. S1) based on our phylogenetic reconstruction and the literature (e.g. Johnson 1980; Mayden 1988; Berendzen et al. 2003, 2008; Haag 2010; Galloway et al. 2011) to infer the demographic history of *P. rubrum* and *P. sintoxia* populations.

Scenario 1 represents the resolved phylogenetic reconstruction generated by SVDquartets; Scenario 2 models the Teays River hypothesis (historical admixture between the Cumberland + Green + Tennessee and Ohio); Scenario 3 represents the last glacial maximum (LGM) hypothesis (historical admixture between the Green + Ohio and Cumberland + Tennessee); Scenario 4 models the Lake Green hypothesis (historical admixture between Cumberland + Ohio + Tennessee and Green); Scenario 5 models the Teays River hypothesis and secondary contact between Interior and Eastern highlands (historical admixture between Cumberland + Green + Tennessee + Ohio, Red + Mississippi + White, and Yazoo + Ouachita); Scenario 6 models the LGM hypothesis and secondary contact between Interior and Eastern highlands; and Scenario 7 models the Lake Green hypothesis and secondary contact between Interior and Eastern highlands (Fig. 2). Visualizations of all models used in DIYABC-RF simulations are presented in Supplementary Fig. S1.

For DIYABC-RF simulations, we removed 12 loci that only had missing data for the Red and Mississippi populations and used more than 10,000 simulated data sets per scenario (75,000 total) to produce posterior distributions with each scenario having equal prior probability. A principal component analysis (PCA) was performed on the summary statistics to evaluate how well simulated data fit observed data from the seven scenarios. We used all simulated data, five noise variables, and 1,000 Random Forest trees to determine the most likely scenario.

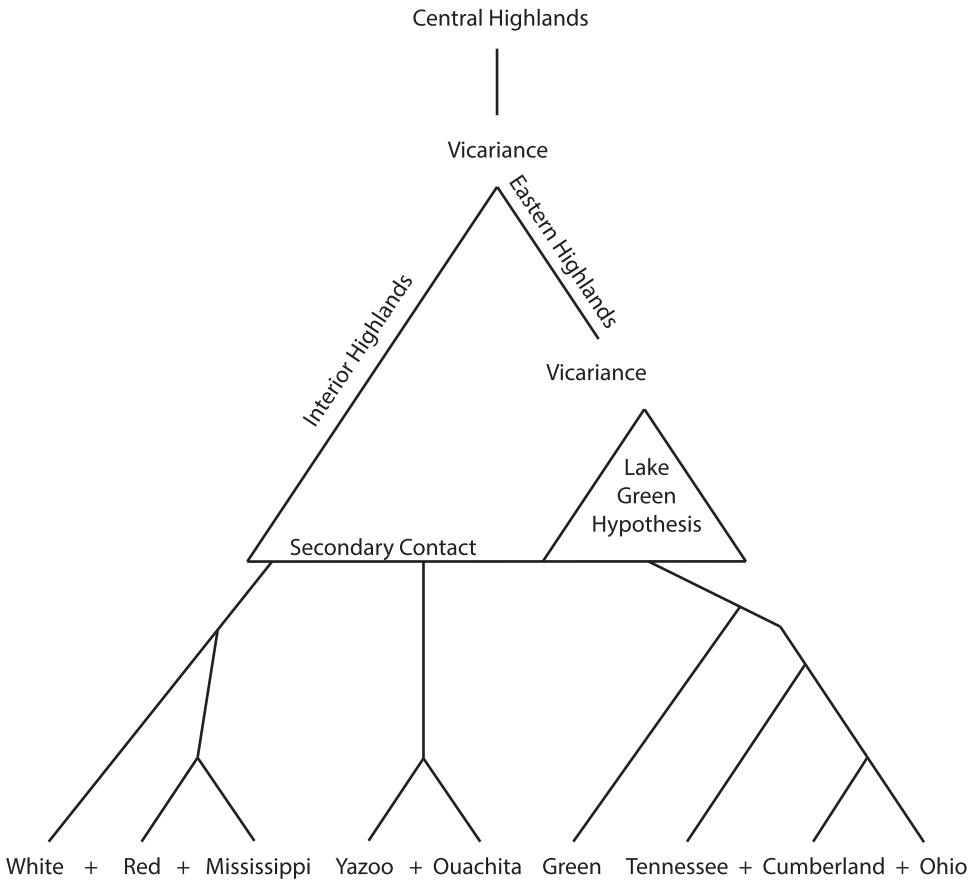

Fig. 2. Graphical representation of the best supported scenario from the DIYABC analysis. All seven scenarios are illustrated in Supplementary Fig. S1.

Tests for hybridization

We tested for evidence of hybridization among individuals from the Eastern Highlands (i.e. Cumberland, Green, Ohio, and Tennessee River drainages) using NewHybrids v 1.1 (Anderson and Thompson 2002). The 200 SNPs with the highest degree of polymorphism for individuals from the Eastern Highlands were used for the analysis. The analysis was conducted for 500,000 sweeps after a 50,000-sweep burn-in. Parent populations were not designated a priori given the observed pattern of genetic variation with respect to geography. Resulting posterior probabilities were used to assess support of individuals being either pure bred (assuming the presence of two parent populations), an F1 hybrid, an F2 hybrid, or a backcross with either parent population.

Results

Molecular data generation

We acquired tissues from 324 specimens originally identified as either *P. rubrum* (n = 99) or *P. sintoxia* (n = 225) to assess genetic differentiation between the two nominal species. From these samples, we successfully generated mtDNA (*COX1* and *ND1*) and nDNA sequences (*ITS1*) for 200 and 106 individuals, respectively. No gaps or stop codons were observed in any of the mtDNA sequences. Each taxon was represented by *COX1* (avg. 657 nucleotides [nt]), *ND1* (avg. 887 nt), and *ITS1* (552 nt). We concatenated sequences from both mtDNA genes (avg. 1,544 nt) for subsequent analyses.

A total of 81,356 polymorphic loci were generated using GBS for 184 individuals. After filtering steps, we retained 176 individuals and 5,366 polymorphic loci. General location, population, and sample sizes for each molecular dataset are presented in Table 1. All DNA data are available from the National Center for Biotechnology Information (*COX1*: OR633009--OR633208; *ND1*: OR635239-OR635438; *ITS1*: OR646411-OR64516; and GBS reads: BioProject IDPRJNA1026624). All metadata associated with specimens and tissues utilized in the study, along with DNA alignments (*COX1*, *ND1*, and *ITS1*) and SNPs are available from Johnson and Smith (2023) on ScienceBase (https://doi.org/10.5066/P9RLSX0Y). Additional details on each specimen used in this study, including museum catalog numbers, collection details, original identifications, and GenBank accession numbers, are available from Johnson and Smith (2023).

Molecular analyses

We were unable to differentiate *P. rubrum* and *P. sintoxia* using nDNA sequences (Fig. 3a), mtDNA sequences (Fig. 3b), or GBS data (Fig. 4a and b), which supports recognition of the nominal taxa as conspecific, or belonging to the same species. Both *ITS1* and mtDNA haplotype networks failed to separate the nominal species and lacked resolution for geographic segregation of individuals based on drainage of capture (Fig. 3a and b). Pairwise genetic distance values are reported in Supplementary Table S1 and ranged from 0.003 to 0.013. Our GBS data generally align with sampling locality according to our PCoA (Fig. 4a), however, we observed patterns of genetic variation within the Eastern Highlands (i.e. Cumberland, Green, Ohio, and Tennessee River basins) that deviated from the IBD model. We did find a significant correlation between genetic and geographic distances ($P < 0.0001$), but the Procrustes analysis showed that samples from the Eastern Highlands, most notably the Green River, deviated from IBD expectations (Supplementary Fig. S2). Our Mantel test supported a significant association ($P < 0.0001$) between all sampled populations and stream distance among sampling localities despite high genetic differentiation at low stream distances (Supplementary Fig. S3), which is driven by comparisons of samples from the Eastern Highlands and congruent with our Procrustes analysis.

Assessments of population structure supported variable numbers of genetic clusters, ranging from K values of 2–9. Pairwise F_{ST} values are reported in Supplementary Table S2 and ranged from 0.03 to 0.4. Using the first eight PCs, DAPC supported K = 9 as the best value (Supplementary Fig. S4) and had an overall assignment score of 91.5%. FastSTRUCTURE indicated K = 6 as the best likelihood and K = 7 as the best K to explain structure in the dataset (Fig. 5). The only difference was the Tennessee basin was a single cluster for K = 6 and split into two distinct clusters for K = 7. Cross-entropy plots from TESS3 supported K = 2 as the best K value but could not differentiate K values of 3–10, likely due to the lack of concordance between observed genetic clustering and geography in the Eastern Highlands (Supplementary Fig. S5). The White River, which represents the Ozark Highlands, and one sample from the Upper Mississippi were distinct from the remainder of the samples for K = 2.

Table 1. Population designations, waterbody of collection, and samples sizes for mitochondrial genes (*COXI* and *ND1*), nuclear DNA sequences (*ITS1*), and genotype-by-sequencing (GBS) datasets.

Pop designations	Waterbody	*COXI*	*ND1*	*ITS1*	GBS
Arkansas	Illinois River	2	2	0	2
Cumberland	Big South Fork	5	5	5	5
	Cumberland River	1	1	1	1
Green	Green River	43	43	29	41
Ohio	Allegheny River	5	5	5	5
	Ohio River	2	2	2	2
	Shenango River	9	9	8	8
Ouachita	Bayou Bartholomew	4	4	0	2
	Ouachita River	5	5	0	0
	Saline River	21	21	6	21
Red	Little River	7	7	5	7
St. Francis	St. Francis River	14	14	6	0
Tennessee	Clinch River	9	9	7	10
	Duck River	6	6	8	6
	Holston River	1	1	0	0
	Tennessee River	7	7	7	7
Upper Mississippi	Chippewa River	6	6	5	6
	Mississippi River	1	1	0	3
White	Black River	12	12	0	12
	Spring River	28	28	2	26
	Strawberry River	4	4	3	3
Yazoo	Big Sunflower River	8	8	7	9
Total		200	200	106	176

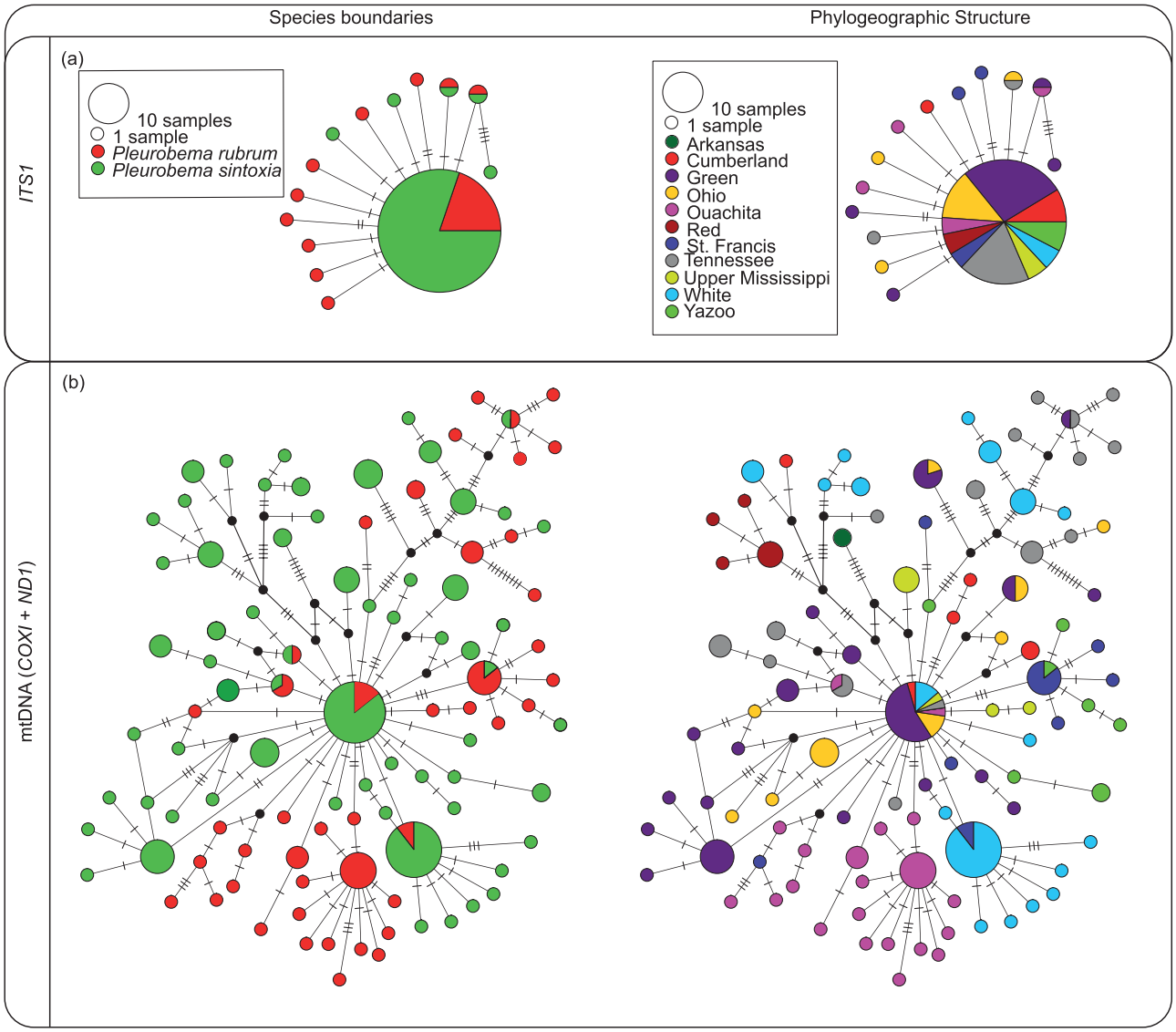

Fig. 3. Haplotype networks used to assess genetic relationships by nominal species (left) and drainage of capture (right) based on a) *ITS1* and b) mtDNA (*COX1* + *ND1*) sequences for *P. rubrum* and *P. sintoxia*. Colors within pie diagrams indicate drainage of capture and size of each circle is proportional to the number of individuals with each haplotype.

Demographic history inferences

The 1,000 Random Forest trees from DIYABC-RF provided the following support values for demographic scenarios (larger values indicate higher support): Scenario 1–80; Scenario 2–106; Scenario 3–91; Scenario 4–140; Scenario 5–182; Scenario 6–201; and Scenario 7–200. DIYABC-RF marginally supported Scenario 6 (posterior probability = 0.493) but had difficulties distinguishing among Scenarios 5–7. Split support among these scenarios was likely driven by the shared support for some level of secondary contact between the Interior and Eastern highlands, but the model could not differentiate scenarios within the Eastern Highlands. The DIYABC-RF supported the LGM and Lake Green hypotheses for the Eastern Highlands (Fig. 2) over the Teays River hypothesis, likely driven by evidence of multiple admixture events in the genetic clusters from the Green River population with other populations in the Eastern Highlands (Fig. 4a and b).

Tests for hybridization

Results from our analyses using NewHybrids strongly supported (all individuals with PP > 0.96) the presence of two parent populations: 1) a portion of individuals from the Cumberland, Green, and Tennessee drainages; and 2) a portion of individuals from the Green and Tennessee drainages. All other individuals were supported as F2 hybrids of the two estimated parent populations (all but one individual PP > 0.97), which included a portion of individuals from the Cumberland, Green, and Tennessee drainages and all the individuals from the Ohio drainage. These findings align with results from our other analyses based on GBS data and support the hypothesis that populations are interbreeding, but remnant ancestral genotypes from historical vicariance have yet to be purged from multiple localities in the Eastern Highlands. The two parent populations were not correlated with initial field-based identifications representing the nominal forms *P. rubrum* and *P. sintoxia*.

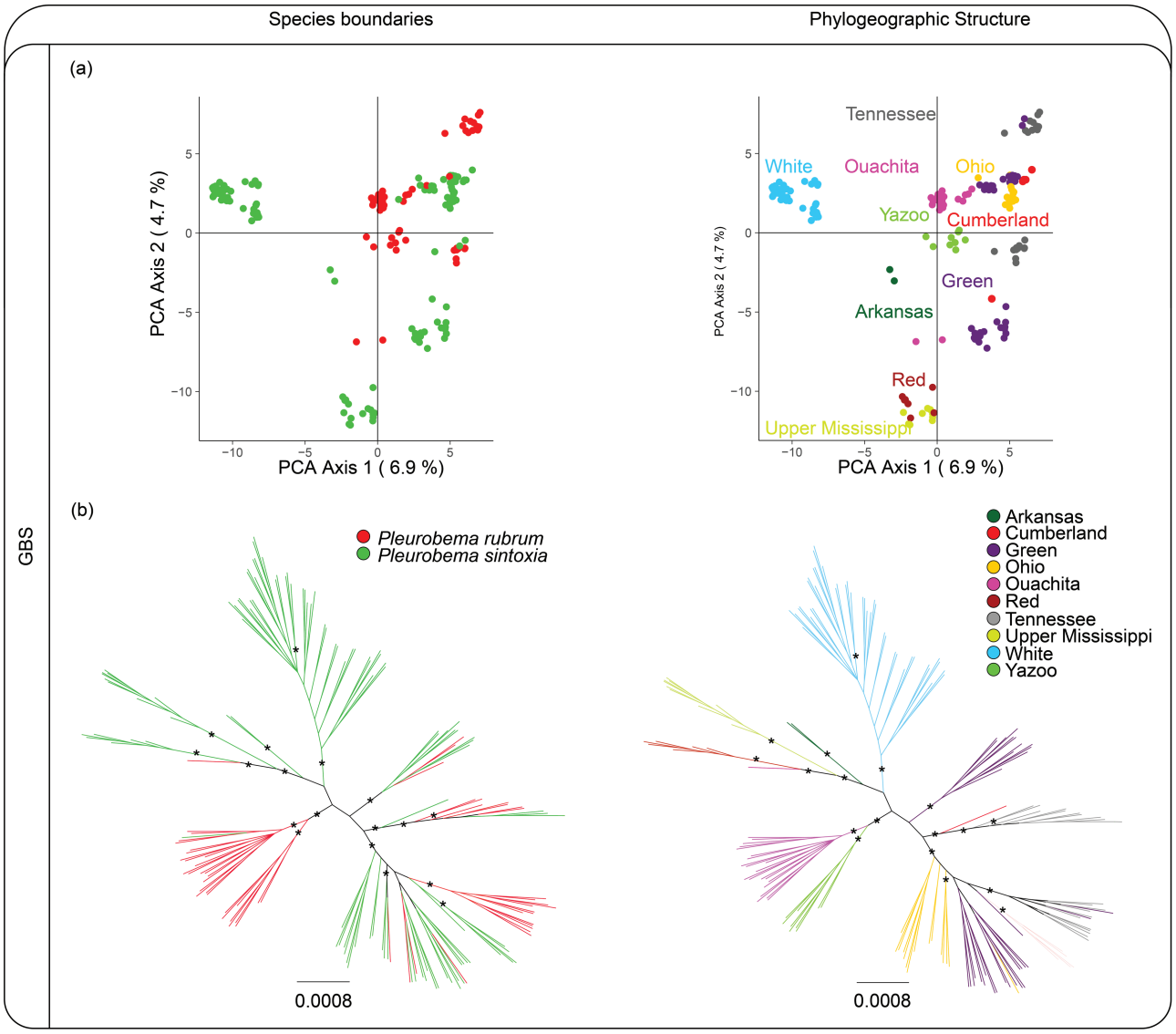

Fig. 4. Depiction of genetic relationships between *P. rubrum* and *P. sintoxia* on our genotype-by-sequencing (GBS) dataset based on a) PCoA and b) unrooted phylogenetic network. Colors correspond to either original field-based identification (left) or drainage of capture (right) and an asterisk denotes a node with UFBS values >90.

Discussion

Our investigation provides the first comprehensive assessment of species boundaries and phylogeography for *P. rubrum* and *P. sintoxia*. The two taxa were not diagnosable using mtDNA sequence, nDNA sequence, or GBS data, and we formally recognize *P. rubrum* as a synonym of *P. sintoxia*. Using the GBS data we collected on both formerly recognized species, we observed patterns of genetic variation that do not follow expectations of IBD, including the retention of three genetic clusters at the same site in the Green River. The three clusters were not unique to the Green River population, each sharing membership with samples from either the Cumberland, Tennessee, or Ohio River basins. This pattern is possibly driven by population expansion from refugia during the LGM, vicariance events following modifications of drainages during Pleistocene glaciation, or infrequent long-distance dispersal events. Our demographic analyses and tests for hybridization provide support for historical vicariance followed by secondary contact. While the timing of these events remains unclear, we hypothesize that genetic variation has been shaped by ancestral vicariance during the formation of Lake Green and subsequent secondary contact after the LGM. Below, we discuss our results in terms of their ecological and evolutionary significance.

Species boundaries in *P. sintoxia* and conservation implications

As in previous studies (e.g. Jones et al. 2015a; Inoue et al. 2018; Olivera-Hyde et al. 2023), *P. rubrum* and *P. sintoxia* were not diagnosable using molecular characters. Our mtDNA and nDNA sequence data showed extensive haplotype sharing among individuals morphologically identified as representative of the two nominal species (Fig. 3a and b). Although mtDNA and nDNA sequence data have been routinely used to test species boundaries in North American freshwater mussels (Pfeiffer et al. 2016; Johnson et al. 2018; Smith et al. 2018, 2021a; Inoue et al. 2020), these markers often lack resolution and show incongruence with existing species-level

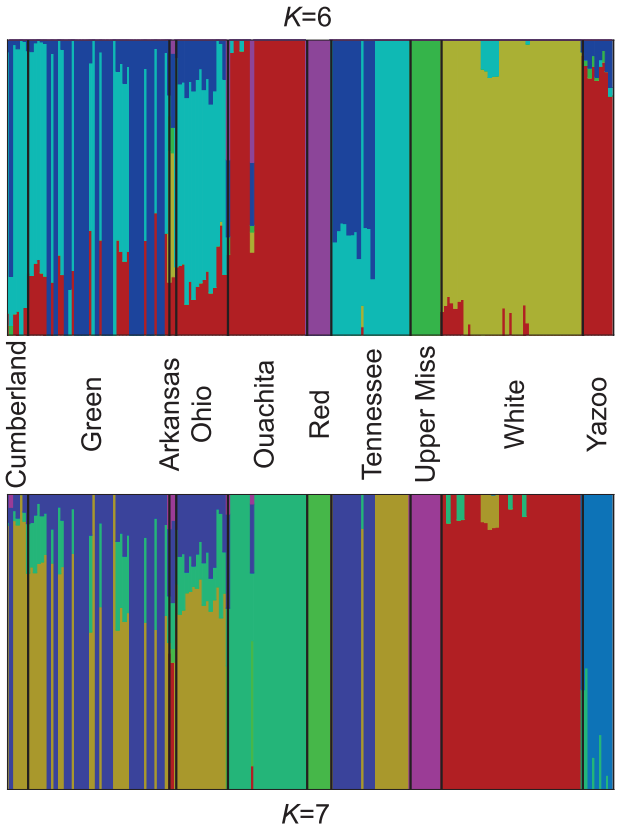

Fig. 5. Genomic cluster assignment based on maximum-likelihood analysis in fastSTRUCTURE at $K = 6$ (top) and $K = 7$ (bottom) based on our genotype-by-sequencing (GBS) dataset for individual *P. sintoxia* samples grouped by drainage of capture.

hypotheses in *Pleurobema* (Campbell et al. 2012; Inoue et al. 2018; Morrison et al. 2021; Olivera-Hyde et al. 2023). To address taxonomic uncertainty and avoid reliance on mitochondrial loci, we used GBS to further investigate species boundaries between *P. rubrum* and *P. sintoxia*. Our GBS data were congruent with findings from mtDNA and nDNA sequence data, indicating a single species with considerable admixture instead of current species hypotheses that recognize two distinct species (Fig. 4a and b). Both *P. sintoxia* and *P. rubrum* were described by Rafinesque in 1820. Following the principle of the first reviser (International Code of Zoological Nomenclature, Article 24.2), we fix the precedence of *P. sintoxia*, which classifies *P. rubrum* as a synonym (ICZN 1999). We selected *P. sintoxia* as the valid name for this species, despite page preference, based on frequency of usage in the literature and broader distribution when compared to *P. rubrum*.

Biodiversity assessments and conservation actions rely on data associated with species names (George and Mayden 2005), making accurate taxonomic classification critical to these efforts. Our findings may help guide pending listing decisions and future conservation and recovery actions for populations formerly included as one or both nominal species. Our data support synonymizing *P. rubrum*, which has been proposed for listing as threatened under the U.S. Endangered Species Act (USFWS 2021), with *P. sintoxia*. The synonymy resolves taxonomic uncertainty and provides new context when assessing the geographic range of *P. sintoxia*. Our assessment of phylogeographic structure and genetic diversity provides a baseline for future recovery actions (if warranted) while giving insights into historical processes shaping the geographic distribution of genetic variation across the Central Highlands region. Our findings add to the growing body of literature that support evaluating morphology-based taxonomic hypotheses using molecular data to inform conservation planning for freshwater mussels (e.g. Pfeiffer et al. 2016; Johnson et al. 2018; Smith et al. 2018, 2019; Inoue et al. 2020; Olivera-Hyde et al. 2023).

Glacial dynamics shape complex demographic history in *P. sintoxia*

Although we found a significant correlation between genetic and geographic distances when analyzing all populations ($P < 0.0001$), genetic variation of *P. sintoxia* individuals from the Eastern Highlands did not meet expectations of IBD (Supplementary Figs. S2 and S3). Therefore, it is unlikely that dispersal or vicariance alone can explain the observed genetic variation in Eastern Highlands drainages. In freshwater taxa, including mussels, glacial refugia are well known to affect the geographic patterns of genetic diversity. Molecular data have been useful in identifying glacial refugia and postglacial dispersal (e.g. Berendzen et al. 2003, 2008; Elderkin et al. 2007, 2008; Hewitt et al. 2019), and we took a similar approach to investigate the effects of drainage evolution and glacial refugia on contemporary patterns of genetic diversity in *P. sintoxia*. Our results suggest that a more complex demographic history, involving ancestral vicariance and secondary contact during post-glaciation dispersal events, best explains the observed genetic differentiation among *P. sintoxia* populations. Our results support that populations from the Interior Highlands (Ozarks and Ouachita highlands) were isolated from the Eastern Highlands (Cumberland, Green, Ohio, Tennessee River drainages), likely during glacial advances in the Pleistocene (Thornbury 1965; Mayden 1988), which aligns with observed patterns seen in other aquatic taxa (e.g. Berendzen et al. 2008).

Our demographic modeling supported a vicariance event isolating the ancestral Green River from other drainages in the Eastern Highlands, which provides an explanation for the observed patterns of contemporary genetic variation in the Green River. During the late Pleistocene, Green River headwaters were isolated by the formation of Lake Green in the lower reach of the Green River mainstem, which has been implicated as a barrier to fish passage (Strange and Burr 1997; Simons 2004; Ray et al. 2006) and likely influenced dispersal of other aquatic organisms. Although the formation of Lake Green provides an explanation for why our results support the early vicariance of the Green River, vicariance alone does not explain observed genetic variation given that individuals collected from the same Green River site were resolved in three genetic clusters, each sharing membership with samples from either the Cumberland, Tennessee, or Ohio river basins (Fig. 4a). Interestingly, a similar pattern of genetic variation has been observed in the Streamlined Chub, *Erimystax dissimilis* (Actinopterygii: Leuciscidae), a fish confirmed to be a larval host for *P. sintoxia* (Culp et al. 2009). *Erimystax dissimilis* mitochondrial haplotypes from the Green River drainage were observed in two divergent clades, one containing haplotypes from all Eastern Highlands drainages sampled and the other exclusive to the Green River (Strange and Burr 1997; Simons 2004). The congruence in patterns of genetic variation in

P. sintoxia and at least one larval host fish suggests a complex demographic history, likely including secondary contact during fish dispersal events, has shaped the genetic makeup of fish and mussel populations in the Green River. Our demographic modeling supported this hypothesis and suggests that secondary contact between the Green River and other drainages in the Eastern Highlands occurred during the evolutionary history of *P. sintoxia*. We hypothesize that secondary contact occurred during contraction of ice sheets after the LGM, albeit we cannot reject the event occurred during other late Pleistocene interglacial periods. This hypothesis is based on the recognition of the Green River basin as an important refugium for freshwater mussels and the hypothesis that the local fauna repopulated the upper Ohio River and its tributaries after glacial events (Johnson 1980). This hypothesis was further supported by our tests for hybridization given all individuals from the Ohio River were identified as putative hybrid origin between putative Green and Tennessee River refugia.

It remains unclear why multiple genetic clusters have been retained in several populations of *P. sintoxia*. The individuals sampled from the Green River were recovered in three genetic clusters that included both morphological forms and samples from other drainages and were supported to interbreed based on hybridization analyses. Similarly, we observed two genetic clusters in the Cumberland and Tennessee River drainages, each of which were most closely related to potential sources from the Green River. We hypothesize the observed patterns of genetic variation are the result of geologic events that created contact zones for different genetic sources to mix during post-glaciation dispersal events. In the Cumberland River drainage, the Big South Fork of the Cumberland River enters the Cumberland River slightly downstream of the hypothesized original location of Cumberland Falls (McGrain 1966). This barrier appears to have allowed the aquatic faunas of the upper and middle Cumberland River to follow independent evolutionary trajectories despite their proximity. Relationships among our genetic samples from the Cumberland drainage are therefore likely explained by this vicariance, with samples from the Big South Fork grouping with a cluster consisting of samples from both the Green and Ohio rivers, while our sample from the lower Cumberland groups most closely with a cluster endemic to the Green River. Our samples from the Tennessee River basin form two distinct clusters, one with samples from the Clinch, Duck, and Green Rivers, and another with only the Clinch and Tennessee rivers. The Clinch and Duck rivers are well known to have complex physiographic histories involving a separation from the Tennessee River and connections to prehistoric drainages in the Eastern Highlands (Mayden 1988). Given the complex demographic history among populations in the region, it is not unexpected that genotypes in both the Clinch and Duck form a cluster with individuals from the Green River drainage.

Population genetic theory (e.g. Hardy 1908) would suggest a single generation of genetic interchange would remove ancestral variation in the populations multiple genetic clusters, but we cannot reject that the observed genetic distinctiveness is functionally significant and may be artifacts of selection for localized conditions. However, the dataset in our study is inadequate to test such hypotheses given the lack of environmental covariates with collection localities and the lack of genomic resources for *Pleurobema* or close relatives. We hypothesize that the variation is selectively neutral, with recent, repeated secondary contact leading to the retention of putative ancestral variation, particularly within populations sampled from the Green River. Further studies using more robust methodologies, such as genome-wide association studies (see Funk et al. 2019), may facilitate our understanding of the functional significance of observed genetic variation in *P. sintoxia* in the Eastern Highlands.

Conclusion

Our findings resolve the longstanding taxonomic uncertainty regarding the conspecific status of *P. rubrum* and *P. sintoxia*, which may have significant conservation implications for North American freshwater mussels. *Pleurobema rubrum* is proposed threatened under the U.S. Endangered Species Act. Our results support the species as a synonym of *P. sintoxia*, which may prompt natural resource managers to reconsider conservation actions for *P. rubrum* and *P. sintoxia*. Our findings also provide another empirical example of how the dynamic geological history of the Eastern Highlands has shaped the demographic history of aquatic species. Future studies using more robust ecological and molecular methodologies may be necessary to better understand the presence of ancestral variation in multiple populations of *P. sintoxia* and its potential adaptive significance.

Supplementary material

Supplementary material is available at *Journal of Heredity* Journal online.

Acknowledgments

We are grateful to Bob Anderson, Brittany Bajo, Bob Butler, Winston Clark, Chris Davidson, Barb Dinkins, Robert Eldridge, Gus Engmann, Hugh Faust, Kyler Hecke, Robert Jones, Jennifer Kagel, Tiffany Leach, Chad Lewis, Chris Lorentz, Brian Mize, Bill Posey, John Spaeth, and Justin Wolbert for assistance with field identifications or collections. We appreciate assistance with specimen curation provided by Elizabeth Black, John Slapcinsky, and Amanda Bemis. Many thanks to Matt Canister and Gaia Meigs-Frend for reviewing the data release associated with this project. The authors wish to thank the Cannatella, Hillis, and Zamudio lab groups at the University of Texas at Austin, Matt Neilson at USGS, and two anonymous reviewers for valuable feedback on earlier drafts of this manuscript.

Funding

Funding for this study was provided by the U.S. Fish and Wildlife Service, U.S. Geological Survey, and the University of Texas at Austin Stengl-Wyer Endowment. The findings and conclusions in the article are those of the authors and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Conflict of interest statement. The authors have no conflict of interest to declare.

Author contributions

NAJ, ARH, JWJ conceived and acquired funding for the study. NAJ, CHS designed the study. NAJ, ARH, JWJ, SAA, GRD, NLE, JTG, JLH, PDH, DWH, TWL, MAM, KRM, CLM, MDW, JDW, CHS contributed samples for the study. NAJ, CEB, CHS performed the laboratory work. NAJ, CHS designed and performed the data analyses. NAJ, CHS wrote the manuscript with all authors contributing edits and discussion.

Benefit-sharing statement

Benefits from this research accrue from the sharing of our datasets and results on public databases as described above. Additionally, all vouchered shell and associated tissues utilized in this study were cataloged and made available at the following public institutions: Arkansas State University Museum of Zoology, Florida Museum, and McClung Museum of Natural History and Culture.

Data accessibility statement

All DNA data are available from the National Center for Biotechnology Information (*COX1*: OR633009--OR633208; *ND1*: OR635239-OR635438; *ITS1*: OR646411-OR64516; and GBS reads: BioProject ID PRJNA1026624). All metadata associated with specimens and tissues utilized in the study, along with DNA alignments (*COX1*, *ND1*, and *ITS1*) and SNPs are available from Johnson and Smith (2023) on ScienceBase (https://doi.org/10.5066/P9RLSX0Y).

References

Anderson EC, Thompson EA. A model-based method for identifying species hybrids using multilocus genetic data. Genetics. 2002:160:1217–1229. doi: 10.1093/genetics/160.3.1217

Avise JC. Phylogeography. Cambridge: Harvard University Press; 2000. doi:10.2307/j.ctv1nzfgj7

Barnhart MC, Haag WR, Roston WN. Adaptations to host infection and larval parasitism in Unionoida. J North Am Benthol Soc. 2008:27:370–394. doi: 10.1899/07-093.1

Berendzen PB, Gamble T, Simons AM. Phylogeography of the bigeye chub *Hybopsis amblops* (Teleostei: Cypriniformes): early Pleistocene diversification and post-glacial range expansion. J Fish Biol. 2008:73:2021–2039. doi: 10.1111/j.1095-8649.2008.02046.x

Berendzen PB, Simons AM, Wood RM. Phylogeography of the northern hogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. J Biogeogr. 2003:30:1139–1152. doi: 10.1046/j.1365-2699.2003.00888.x

Bernatchez L, Wilson CC. Comparative phylogeography of Nearctic and Palearctic fishes. Mol Ecol. 1998:7:431–452. doi: 10.1046/j.1365-294x.1998.00319.x

Breton S, Stewart DT, Shepardson S, Trdan RJ, Bogan AE, Chapman EG, Ruminas AJ, Piontkivska H, Hoeh WR. Novel protein genes in animal mtDNA: a new sex determination system in freshwater mussels (Bivalvia: Unionoida)? Mol Biol Evol. 2011:28:1645–1659. doi: 10.1093/molbev/msq345

Campbell DC, Lydeard C. The genera of Pleurobemini (Bivalvia: Unionidae: Ambleminae). Am Malacol Bull. 2012:30:19–38. doi: 10.4003/006.030.0102

Campbell DC, Serb JM, Buhay JE, Roe KJ, Minton RL, Lydeard C. Phylogeny of North American amblemines (Bivalvia, Unionoida): prodigious polyphyly proves pervasive across genera. Invertebr Biol. 2005:124:131–164. doi: 10.1111/j.1744-7410.2005.00015.x

Caye K, Deist TM, Martins H, Michel O, François O. TESS3: fast inference of spatial population structure and genome scans for selection. Mol Ecol Resour. 2016:16:540–548. doi: 10.1111/1755-0998.12471

Chifman J, Kubatko L. Quartet inference from SNP data under the coalescent model. Bioinformatics. 2014:30:3317–3324. doi: 10.1093/bioinformatics/btu530

Collin FD, Durif G, Raynal L, Lombaert E, Gautier M, Vitalis R, Marin JM, Estoup A. Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. Mol Ecol Resour. 2021:21:2598–2613. doi: 10.1111/1755-0998.13413

Courtois B, Audebert A, Dardou A, Roques S, Ghneim-Herrera T, Droc G, Frouin J, Rouan L, Gozé E, Kilian A, et al. Genome-wide association mapping of root traits in a japonica rice panel. PLoS One. 2013:8:1–18. doi: 10.1371/journal.pone.0078037

Crandall KA, Buhay JE. Global diversity of crayfish (Astacidae, Cambaridae, and Parastacidae—Decapoda) in freshwater. Hydrobiologia. 2008:595:295–301. doi:10.1007/s10750-007-9120-3

Cruz VMV, Kilian A, Dierig DA. Development of DArT marker platforms and genetic diversity assessment of the U.S. Collection of the new oilseed crop *Lesquerella* and related species. PLoS One. 2013:8:e64062. doi: 10.1371/journal.pone.0064062

Culp JJ, Shepard AC, McGregor MA. Fish hosts and conglutinates of the pyramid pigtoe (*Pleurobema rubrum*). Southeast Nat. 2009:8:19–22. doi: 10.1656/058.008.0102

Ehlers J, Gibbard PL, Hughes PD, editors. Quaternary glaciations-extent and chronology: a closer look. Vol. 15. Amsterdam: Elsevier; 2011.

Elderkin CL, Christian AD, Metcalfe-Smith JL, Berg DJ. Population genetics and phylogeography of freshwater mussels in North America, *Elliptio dilatata* and *Actinonaias ligamentina* (Bivalvia: Unionidae). Mol Ecol. 2008:17:2149–2163. doi: 10.1111/j.1365-294X.2008.03745.x

Elderkin CL, Christian AD, Vaughn CC, Metcalfe-Smith JL, Berg DJ. Population genetics of the freshwater mussel, *Amblema plicata* (Say 1817) (Bivalvia: Unionidae): evidence of high dispersal and post-glacial colonization. Conserv Genet. 2007:8:355–372. doi: 10.1007/s10592-006-9175-0

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One. 2011:6:e19379–e19310. doi: 10.1371/journal.pone.0019379

Etnier DA, Starnes WC. The fishes of Tennessee. Knoxville: University of Tennessee Press; 1993.

Funk WC, Forester BR, Converse SJ, Darst C, Morey S. Improving conservation policy with genomics: a guide to integrating adaptive potential into US Endangered Species Act decisions for conservation practitioners and geneticists. Conserv Genet. 2019:20:115–134. doi: 10.1007/s10592-018-1096-1

Galloway WE. Depositional evolution of the Gulf of Mexico Sedimentary Basin. In: Miall AD, editors. Sedimentary Basins of the World. Vol. 5. Amsterdam: Elsevier; 2008. doi:10.1016/S1874-5997(08)00015-4

Galloway WE, Whiteaker TL, Ganey-Curry P. History of Cenozoic North American drainage basin evolution, sediment yield, and accumulation in the Gulf of Mexico basin. Geosphere. 2011:7:938–973. doi: 10.1130/GES00647.1

George AL, Mayden RL. Species concepts and the endangered species act: how a valid biological definition of species enhances the legal protection of biodiversity. Nat Res J. 2005:45:369–407.

Georges A, Gruber B, Pauly GB, White D, Adams M, Young MJ, Kilian A, Zhang X, Shaffer HB, Unmack PJ. Genomewide SNP markers breathe new life into phylogeography and species delimitation for the problematic short-necked turtles (Chelidae: *Emydura*) of eastern Australia. Mol Ecol. 2018:27:5195–5213. doi: 10.1111/mec.14925

Graf DL, Cummings KS. A "big data" approach to global freshwater mussel diversity (Bivalvia: Unionoida), with an updated checklist of

genera and species. J Molluscan Stud. 2021:87:1–36. doi: 10.1093/mollus/eyaa034
Haag WR. A hierarchical classification of freshwater mussel diversity in North America. J Biogeogr. 2010:37:12–26. doi: 10.1111/j.1365-2699.2009.02191.x
Haag WR. North American freshwater mussels: natural history, ecology, and conservation. New York: Cambridge University Press; 2012. doi:10.1017/CBO9781139048217
Haag WR, Cicerello RR. A distributional atlas of the freshwater mussels of Kentucky. Scientific and Technical Series 8. Frankfort: Kentucky State Nature Preserves Commission; 2016.
Haag WR, Williams JD. Biodiversity on the brink: an assessment of conservation strategies for North American freshwater mussels. Hydrobiologia. 2014:735:45–60. doi: 10.1007/s10750-013-1524-7
Hardy GH. Mendelian proportions in a mixed population. Science. 1908:28:49–50.
Henley WF, Grobler PJ, Neves RJ. Non-invasive method to obtain DNA from freshwater mussels (Bivalvia: Unionidae). J Shellfish Res. 2006:25:975–977. doi: 10.2983/0730-8000(2006)25[975:NMTODF]2.0.CO;2
Hewitt TL, Haponski AE, Foighil D. Evolution of diverse host infection mechanisms delineates an adaptive radiation of lampsiline freshwater mussels centered on their larval ecology. PeerJ. 2021:9:e12287.
Hewitt TL, Woolnough DA, Zanatta DT. Population genetic analyses of *Lampsilis cardium* (Bivalvia: Unionida) reveal multiple post-glacial colonization routes into the Great Lakes drainage. Am Malacol Bull. 2019:37:21–34. doi: 10.4003/006.037.0103
ICZN (International Commission on Zoological Nomenclature). International Code of Zoological Nomenclature. 4th edn. London: International Trust for Zoological Nomenclature; 1999. p. xxix + 306.
Inoue K, Cummings KS, Tiemann JS, Miller TD, Johnson NA, Smith CH, Randklev CR. A new species of freshwater mussel in the genus *Popenaias* Frierson, 1927, from the Gulf coastal rivers of central Mexico (Bivalvia: Unionida: Unionidae) with comments on the genus. Zootaxa. 2020:4816:457–490. doi: 10.11646/zootaxa.4816.4.3
Inoue K, Hayes DM, Harris JL, Johnson NA, Morrison CL, Eackles MS, King TL, Jones JW, Hallerman EM, Christian AD, et al. The Pleurobemini (Bivalvia: Unionida) revisited: Molecular species delineation using a mitochondrial DNA gene reveals multiple conspecifics and undescribed species. Invertebr Syst. 2018:32:689–702. doi: 10.1071/IS17059
Inoue K, Monroe EM, Elderkin CL, Berg DJ. Phylogeographic and population genetic analyses reveal Pleistocene isolation followed by high gene flow in a wide ranging, but endangered, freshwater mussel. Heredity. 2014:112:282–290. doi: 10.1038/hdy.2013.104
Johnson NA, Beaver CE, Kiser AH, Duplessis M, Wagner M, Ellwanger R, Robertson CR, Kinney SD, Gregory BB, Wolverton S, et al. Molecular data validate the historical and contemporary distribution of *Pleurobema riddellii* (Bivalvia: Unionidae) and help guide conservation and recovery efforts. Endanger Species Res. 2023:52:1–15. doi: 10.3354/esr01266
Johnson NA, Smith CH. Molecular data used to test species boundaries, characterize phylogeographic patterns of genetic diversity, and guide the Endangered Species Act listing decision for a North American freshwater mussel species complex. U.S. Geological Survey data release. 2023. doi:10.5066/P9RLSX0Y
Johnson NA, Smith CH, Pfeiffer JM, Randklev CR, Williams JD, Austin JD. Integrative taxonomy resolves taxonomic uncertainty for freshwater mussels being considered for protection under the US Endangered Species Act. Sci Rep. 2018:8:15892. doi: 10.1038/s41598-018-33806-z
Johnson RI. Zoogeography of North American Unionacea (Mollusca: Bivalvia) north of the maximum Pleistocene glaciation. Bull Mus Comp Zool. 1980:149:77–189.
Johnson RI, Baker HB. The types of Unionacea (Mollusca: Bivalvia) in the Academy of Natural Sciences of Philadelphia. Proc Acad Nat Sci Phila. 1973:125:145–186.
Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008:24:1403–1405. doi: 10.1093/bioinformatics/btn129
Jombart T, Ahmed I. adegenet 13-1: new tools for the analysis of genome-wide SNP data. Bioinformatics. 2011:27:3070–3071. doi: 10.1093/bioinformatics/btr521
Jones JW, Johnson NA, Grobler P, Schilling D, Neves RJ, Hallerman EM. Endangered rough *Pigtoe pearlymussel*: assessment of phylogenetic status and genetic differentiation of two disjunct populations. J Fish Wildl Manag. 2015a:6:338–349. doi: 10.3996/052013-JFWM-036
Jones JW, Neves RJ, Hallerman EM. Historical demography of freshwater mussels (Bivalvia: Unionidae): genetic evidence for population expansion and contraction during the late Pleistocene and Holocene. Biol J Linn Soc. 2015b:114:376–397.
Jones RL, Slack WT, Hartfield PD. The freshwater mussels (Mollusca: Bivalvia: Unionidae) of Mississippi. Southeast Nat. 2005:4:77–92. doi: 10.1656/1528-7092(2005)004[0077:TFMMBU]2.0.CO;2
Jones RL, Wagner MD, Slack WT, Peyton S, Hartfield P. Guide to the identification and distribution of freshwater mussels (Bivalvia: Unionidae) in Mississippi. 2nd ed. Jackson: Mississippi Department of Wildlife, Fisheries, and Parks; 2021.
Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:30:772–780. doi: 10.1093/molbev/mst010
Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012:28:1647–1649. doi: 10.1093/bioinformatics/bts199
Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, Caig V, Heller-Uszynska K, Jaccoud D, Hopper C, et al. Diversity arrays technology: a generic genome profiling technology on open platforms. In: Pompanon F, Bonin A, editors. Data production and analysis in population genomics: methods and molecular biology. New Jersey: Humana Press; 2012. p. 67–89. doi:10.1007/978-1-61779-870-2_5
Kinziger AP, Wood RM, Welsh SA. Systematics of *Etheostoma tippecanoe* and *Etheostoma denoncourti* (Perciformes: Percidae). Copeia. 2001:2001:235–239. doi: 10.1643/0045-8511(2001)001[0235:SOETAE]2.0.CO;2
Leigh JW, Bryant D. POPART: full-feature software for haplotype network construction. Methods Ecol Evol. 2015:6:1110–1116. doi: 10.1111/2041-210X.12410
Lopes-Lima M, Emre M, Kebapçı U, Hülya S, Yanık T, Mirzajani A, et al. Molecular phylogenetics and evolution diversity, biogeography, evolutionary relationships, and conservation of Eastern Mediterranean freshwater mussels (Bivalvia: Unionidae). Mol Phylogenet Evol. 2021a:163:107261. doi: 10.1016/j.ympev.2021.107261
Lopes-Lima M, Riccardi N, Urbanska M, Köhler F, Vinarski M, Bogan AE, Sousa R. Major shortfalls impairing knowledge and conservation of freshwater molluscs. Hydrobiologia. 2021b:848:2831–2867. doi: 10.1007/s10750-021-04622-w
Lovelace R, Ellison R. stplanr: a package for transport planning. R J. 2018:10:7–23. doi: 10.32614/RJ-2018-053
Lydeard C, Mayden RL. A diverse and endangered aquatic ecosystem of the southeast United States. Conserv Biol. 1995:9:800–805.
Maddidon WP, Maddison DR. Mesquite: a modular system for evolutionary analysis. Version 3.81 [accessed 2023 April]. http://www.mesquiteproject.org.
Mayden RL. Biogeography of Ouachita Highland fishes. Southwest Nat. 1985:30:195–211.
Mayden RL. Vicariance biogeography, parsimony, and evolution in North American freshwater fishes. Syst Zool. 1988:37:329–355.
McGrain P. Geology of the Cumberland Falls State Park area. Lexington: Kentucky Geological Survey, Series 10, Special Publication, 11; 1966. p. 1–33.

Mijangos JL, Gruber B, Berry O, Pacioni C, Georges A. dartR v2: an accessible genetic analysis platform for conservation, ecology and agriculture. Methods Ecol Evol. 2022:13:2150–2158. doi: 10.1111/2041-210X.13918

Morrison CL, Johnson NA, Jones JW, Eackles MS, Aunins AW, Fitzgerald DB, Hallerman EM, King TL. Genetic and morphological characterization of the freshwater mussel clubshell species complex (*Pleurobema clava* and *Pleurobema oviforme*) to inform conservation planning. Ecol Evol. 2021:11:15325–15350. doi: 10.1002/ece3.8219

Near TJ, Page LM, Mayden RL. Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. Mol Ecol. 2001:10:2235–2240. doi: 10.1046/j.1365-294X.2001.01362.x

Neemuchwala S, Johnson NA, Pfeiffer JM, Lopes-Lima M, Gomes-dos-Santos A, Froufe E, Hillis DM, Smith CH. Coevolution with host fishes shapes parasitic life histories in a group of freshwater mussels (Unionidae: Quadrulini). Bull Soc Syst Biol. 2023:2:1–25. doi: 10.18061/bssb.v2i1.8998

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, OHara RB, Simpson GL, Solymos P, et al. vegan: Community ecology package. R package version 2.5-7. 2022.

Olivera-Hyde M, Jones JW, Hallerman EM. Phylogenetic assessment of endangered and look-alike Pigtoe species in a freshwater mussel diversity hotspot. Ecol Evol. 2023:13:e9717. doi: 10.1002/ece3.9717

Ortmann AE. A monograph of the Najades of Pennsylvania. In: Memoirs of the Carnegie Museum. Vol. 4. Pittsburgh: The authority of the Board of Trustees of the Carnegie Institute, 1901–1936. 1911. p. 279–347. doi:10.5962/p.48330

Ortmann AE. Correlation of shape and station in fresh-water mussels (Naiades). Proc Am Philos Soc. 1920:59:268–312. doi: 10.2307/984426

Ortmann AE. The naiad-fauna of Duck River in Tennessee. Am Midl Nat. 1924:9:18–62.

Ortmann AE. The naiad-fauna of the Tennessee River system below Walden Gorge. Am Midl Nat. 1925:9:321–372.

Parmalee P, Bogan A. The freshwater mussels of Tennessee. Knoxville: University of Tennessee Press; 1998.

Pebesma E. Simple features for R: standardized support for spatial vector data Version 08. R J. 2018:10:439–446.

Perkins MA, Johnson NA, Gangloff MM. Molecular systematics of the critically-endangered North American spinymussels (Unionidae: *Elliptio* and *Pleurobema*) and description of *Parvaspina* gen nov. Conserv Genet. 2017:18:745–757. doi: 10.1007/s10592-017-0924-z

Pfeiffer JM, Johnson NA, Randklev CR, Howells RG, Williams JD. Generic reclassification and species boundaries in the rediscovered freshwater mussel "*Quadrula*" *mitchelli* (Simpson in Dall, 1896). Conserv Genet. 2016:17:279–292. doi: 10.1007/s10592-015-0780-7

R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022 [accessed 2023 April]. https://www.R-project.org/

Raj A, Stephens M, Pritchard JK. FastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics. 2014:197:573–589. doi: 10.1534/genetics.114.164350

Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol G, Edwards D, McCully M, et al. Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. PLoS One. 2014:9:e101673. doi: 10.1371/journal.pone.0101673

Ray JM, Wood RM, Simons AM. Phylogeography and post-glacial colonization patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). J Biogeogr. 2006:33:1550–1558. doi: 10.1111/j.1365-2699.2006.01540.x

Robison HW. Zoogeographic implications of the Mississippi River basin, pp. 267-283. In: Hocutt CH, Wiley EO, editors. The zoogeography of North American Freshwater Fishes. New York: John Wiley and Sons, Inc; 1986. p. 866.

Sansaloni C, Petroli C, Jaccoud D, Carling J, Detering F, Grattapaglia D, Kilian A. Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of *Eucalyptus*. BMC Proc. 2011:5:1–2. doi: 10.1186/1753-6561-5-s7-p54

Shea CP, Peterson JT, Wisniewski JM, Johnson NA. Misidentification of freshwater mussel species (Bivalvia: Unionidae): contributing factors, management implications, and potential solutions. J North Am Benthol Soc. 2011:30:446–458. doi: 10.1899/10-073.1

Simons AM. Phylogenetic relationships in the genus *Erimystax* (Actinopterygii: Cyprinidae) based on the cytochrome b gene. Copeia. 2004:2004:351–356. doi: 10.1643/CG-02-289R2

Slatkin M. Gene flow in natural populations. Annu Rev Ecol Syst. 1985:16:393–430. doi: 10.1146/annurev.ecolsys.16.1.393

Smith CH, Johnson NA, Inoue K, Doyle RD, Randklev CR. Integrative taxonomy reveals a new species of freshwater mussel, *Potamilus streckersoni* sp nov (Bivalvia: Unionidae): implications for conservation and management. Syst Biodivers. 2019:17:331–348. doi: 10.1080/14772000.2019.1607615

Smith CH, Johnson NA, Pfeiffer JM, Gangloff MM. Molecular and morphological data reveal non-monophyly and speciation in imperiled freshwater mussels (*Anodontoides* and *Strophitus*). Mol Phylogenet Evol. 2018:119:50–62. doi: 10.1016/j.ympev.2017.10.018

Smith CH, Johnson NA, Havlik K, Doyle RD, Randklev CR. Resolving species boundaries in the critically imperiled freshwater mussel species, *Fusconaia mitchelli* (Bivalvia: Unionidae). J Zool Syst Evol Res. 2021a:59:60–77. doi: 10.1111/jzs.12412

Smith CH, Johnson NA, Robertson CR, Doyle RD, Randklev CR. Establishing conservation units to promote recovery of two threatened freshwater mussel species (Bivalvia: Unionida: *Potamilus*). Ecol Evol. 2021b:11:11102–11122. doi: 10.1002/ece3.7897

Smith CH, Pfeiffer JM, Johnson NA. Comparative phylogenomics reveal complex evolution of life history strategies in a clade of bivalves with parasitic larvae (Bivalvia: Unionoida: Ambleminae). Cladistics. 2020:36:505–520. doi: 10.1111/cla.12423

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS. Comparative phylogeography of unglaciated eastern North America. Mol Ecol. 2006:15:4261–4293. doi: 10.1111/j.1365-294X.2006.03061.x

Starnes WC, Etnier DA. Drainage evolution and fish biogeography of the Tennessee and Cumberland Rivers Drainage realm. In:Hocutt CH, Wiley EO, editors. The Zoogeography of North American freshwater fishes. New York: Wiley, 1986. p. 325–361.

Strange RM, Burr BM. Intraspecific phylogeography of North American highland fishes: a test of the Pleistocene vicariance hypothesis. Evolution. 1997:51:885–897. doi: 10.1111/j.1558-5646.1997.tb03670.x

Strayer DL, Dudgeon D. Freshwater biodiversity conservation: recent progress and future challenges. J North Am Benthol Soc. 2010:29:344–358. doi: 10.1899/08-171.1

Switzer JF, Wood RM. Molecular systematics and historical biogeography of the Missouri saddled darter *Etheostoma tetrazonum* (Actinopterygii: Percidae). Copeia. 2002:2002:450–455. doi: 10.1643/0045-8511(2002)002[0450:MSAHBO]2.0.CO;2

Swofford D. PAUP*. Phylogenetic analysis using Parsimony (*and other methods). Version 4.0b10. 2002. doi: 10.1111/j.0014-3820.2002.tb00191.x

Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021:38:3022–3027. doi: 10.1093/molbev/msab120

Thornbury WD. Regional geomorphology of the United States. New York: John Wiley and Sons; 1965.

U.S. Fish and Wildlife Service (USFWS). Endangered and threatened wildlife and plants; threatened species status with section 4(d) rule for Pyramid Pigtoe. Fed Regist. 2021:75:56928–56935. doi: 10.1016/0196-335x(80)90058-8

Vanatta EG. Rafinesque's types of Unio. Proc Acad Nat Sci Phila. 1915:67:549–559.

Vidrine MF. The historical distributions of freshwater mussels in Louisiana. Louisiana: Gail Q. Vidrine Collectables; 1993.

Wang C, Szpiech ZA, Degnan JH, Jakobsson M, Pemberton TJ, Hardy JA, Singleton AB, Rosenberg NA. Comparing spatial maps of human population-genetic variation using procrustes analysis. Stat Appl Genet Mol Biol. 2010:9:1–20. doi: 10.2202/1544-6115.1493

Watters GT, Hoggarth MA, Stansbery DH. The freshwater mussels of Ohio. Columbus: The Ohio State University Press; 2009. doi: 10.5860/choice.47-5655

Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A. Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proc Natl Acad Sci USA. 2004:101:9915–9920. doi: 10.1073/pnas.0401076101

Williams JD, Bogan AE, Butler RS, Cummings KS, Garner JT, Harris JL, Johnson NA, Watters GT. A revised list of the freshwater mussels (Mollusca: Bivalvia: Unionida) of the United States and Canada. Freshw Mollusk Biol Conserv. 2017:20:33–58. doi: 10.31931/fmbc.v20i2.2017.33-58

Williams JD, Bogan AE, Garner JT. Freshwater mussels of Alabama and the Mobile Basin in Georgia, Mississippi and Tennessee. Tuscaloosa: University of Alabama Press; 2008.

Williams JD, Warren ML, Cummings KS, Harris JL, Neves RJ. Conservation status of freshwater mussels of the United States and Canada. Fisheries. 1993:18:6–22. doi: 10.1577/1548-8446(1993)018<0006:csofmo>2.0.co;2

Wright S. Isolation by distance. Genetics. 1943:28:114–138.