

Phylogeography and population genetics of a headwater-stream adapted crayfish,  
*Cambarus pristinus* (Decapoda: Cambaridae), from the Cumberland Plateau in Tennessee

A thesis

Presented to

The College of Graduate Studies

Austin Peay State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science in Biology

Brooke A. Grubb

December, 2019

### Statement of Permission to Use

In presenting this thesis in partial fulfillment of the requirements for the Master of Science in Biology at Austin Peay State University, I agree that the library shall make it available to borrowers under the rules of the library. Brief quotations from this field study are allowable without special permission, provided that accurate acknowledgement of the source is made. Permissions for extensive quotations or reproduction of this field study may be granted by my major professor, or in his/her absence, by the Head of the Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Brooke A. Grubb

Date:

---

---

For Rose Mier  
She introduced me to the  
wonderful world of crayfish.

## **ACKNOWLEDGMENTS**

First, I want to thank Dr. Rebecca Blanton Johansen for all her help, guidance, and feedback during my master's work. Without Rebecca focusing my enthusiasm, I likely would have followed every 'shiny object' and my master's work would've been a much longer process. I also want to thank her for all the invaluable advice related to pursuing a PhD and being available to answer any and all questions I might have had and thank her for allowing me to pursue my Ph.D. while finishing my master's work, despite the difficulties of working long distance.

I would also like to thank Dr. John Johansen. His experience with crayfishes in Tennessee has been an invaluable source of information. I also want to thank him for always helping in the field. His field commentary always brought group morale up and his knowledge of the area proved invaluable.

I want to thank Rebecca and John both for letting me stay with them my last few months in Tennessee and for constantly being sounding boards for my thoughts.

I also want to thank Dr. Mollie Cashner who has provided invaluable advice when I was having trouble with lab work and for asking questions which would make me look at my work from a different angle. I would like to thank her and Dr. Steve Hamilton for serving on my committee and providing feedback on my thesis.

My day to day life would have been bland without all the amazing friends I've made in the graduate program. You all provided an amazing amount of support over the last couple of years. You pitched in with field work, lab work, watched my pets when needed, listened to my grievances, climbed with me, and ultimately were there whenever I needed you. May the pun board and bread phylogeny live on in infamy! Thank you:

Nastasia Disotell, Jacob Brumley, Ben Wilson, Brooke Washburn, Josh Stonecipher, Allison Kirby, Bailey Harris, John and Rebecca Johansen, Claire Ciafre, Mollie Cashner, Lyranda Thiem, Thomas Murphy, Sandra Bojic, Beau Rapier, Libnis Sanchez, Tammy St. James, Chelsea Blumbergs, Mariah Schlis-Elias, Caroline Wilford.

I would also like to thank Chris Skelton, Carl Williams, and Jeff Simmons for their help in finding and capturing crayfish. Your field experience allowed us to find individuals in an area that we had been unsuccessful in. I would like to thank Martha Lewis for allowing us to collect on her property and Matt Brown for aiding me in getting permits required for collecting at Fall Creek Falls State Park. I would also like to thank the Tennessee Wildlife Resource Agency and Austin Peay's Center of Excellence for Field Biology for funding which made this project possible

## ABSTRACT

BROOKE A. GRUBB. Phylogeography and population genetics of a headwater-stream adapted crayfish, *Cambarus pristinus* (Decapoda: Cambaridae), from the Cumberland Plateau in Tennessee. (Under the direction of DR. REBECCA BLANTON JOHANSEN)

Habitat loss and fragmentation represent significant threats to North American crayfish diversity. Assessments of genetic diversity within and among populations of imperiled species can provide a baseline for determining the relative impacts of contemporary anthropogenic threats such as habitat fragmentation to population connectivity, as well as aid in identifying historical factors that contribute to population structure. *Cambarus pristinus*, a species of conservation concern and a headwater-stream adapted crayfish endemic to the Cumberland Plateau in Tennessee, exhibits a disjunct distribution within 4<sup>th</sup> order or lower tributaries of the Caney Fork and Big Brush Creek systems and is comprised of two morphologically distinct forms, the nominal Caney Fork form (Caney Fork system) and the Sequatchie form (Caney Fork and Big Brush Creek systems). Habitat degradation from activities such as silviculture and strip mining have been observed throughout the Cumberland Plateau and *Cambarus pristinus* has experienced recent local extirpations. Our objectives were to examine variation in mitochondrial DNA (COI) and nuclear alleles (microsatellites) to provide estimates of phylogeographic relationships and contemporary levels of genetic structure, respectively, within *C. pristinus*. We predicted that changes in stream hydrology, physiographic regions, and drainage divides would contribute to long-standing isolation among populations separated by these features. We also expected recent anthropogenic disturbances and population loss to have further impacted population connectivity and

lead to reduced genetic diversity and increased population isolation relative to historic levels. Assessment of variation in mitochondrial haplotypes and phylogeographic relationships found low haplotype divergences and broadly shared haplotypes within each morphological form, implying that gene flow was maintained among populations within a form at some level historically, and that geographic features and natural instream barriers did not prevent dispersal. Each form had a unique set of haplotypes and was recovered as a separate, divergent clade indicating the two forms represent distinct genetic lineages and supporting recognition of the Sequatchie form as a distinct taxon. For the Caney Fork form of *C. pristinus* our microsatellite data recovered a high degree of population isolation and support for the occurrence of six isolated populations. We also recovered several low genetic diversity metrics within each cluster and for the Caney Fork form overall. This suggests that the Caney Fork form of *C. pristinus* has a reduced adaptive potential and that historic connectivity has been lost under anthropogenic disturbance. We suggest that *Cambarus pristinus* warrants continued state protection and that future population genetic monitoring be implemented.

## TABLE OF CONTENTS

CHAPTER	
PAGE	
<b>I. Introduction.....</b>	<b>1</b>
<b>II. Methods.....</b>	<b>7</b>
Sample collections.....	7
DNA extractions.....	8
Mitochondrial DNA data collection and analyses.....	8
Microsatellite genotype data collection.....	9
Microsatellite marker validation.....	10
Spatial genetic structure.....	11
Genetic diversity estimation.....	12
<b>III. Results.....</b>	<b>13</b>
Sample collections.....	13
Mitochondrial DNA data collection and analyses.....	14
Microsatellite marker validation.....	15
Spatial genetic structure.....	15
Genetic diversity estimates.....	17
<b>IV. Discussion.....</b>	<b>18</b>
Phylogeography of <i>Cambarus pristinus</i> .....	19
Contemporary genetic diversity.....	23
Conservation implications for <i>Cambarus pristinus</i> .....	28



TABLES.....	31
FIGURES.....	34
LITERATURE CITED.....	42
APPENDIX A.....	58

## List of Tables

**Table 1.** Locality information and number of individuals of *Cambarus pristinus* captured.

A “+” denotes new localities identified and sampled for *Cambarus pristinus*. Site ID corresponds to those used in all other tables and figures.....31

**Table 2.** Genetic diversity measures based on 19 microsatellite loci for 6 distinct population clusters examined for the Caney Fork form of *Cambarus pristinus*: mean alleles per locus ( $N_a$ ), allelic richness (AR), private allelic richness (PAR), observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity, inbreeding coefficient ( $F_{IS}$ ), p-values for Bottleneck, p-values for Hardy-Weinberg equilibrium (HWE), and percentage of polymorphic loci. Values in bold were significant ( $p < 0.05$  for  $F_{IS}$  and Bottleneck, and  $p < 0.003$  for HWE after Bonferroni correction); n is the number examined per distinct population cluster...32

**Table 3.** Pairwise measures of genetic and geographic distances used for the IBD analysis of the Caney Fork form of *Cambarus pristinus* based on 19 microsatellite loci (Figure 4). Site ID corresponds to those used in Figure 1 and defined in Table 1. a) Pairwise  $F_{ST}$  values (above the diagonal) and pairwise Jost’s D values (below the diagonal). Values in bold are significant at  $p < 0.05$  and values with an asterisk (\*) are significant following Bonferroni correction ( $p < 0.0018$ ). b) Pairwise river distance (above the diagonal) and pairwise Euclidean distance (below the diagonal) in kilometers.....33

**Table S1.** Genetic diversity metrics summarized from the literature for several imperiled and non-imperiled crayfish and compared to those for *Cambarus pristinus* generated herein. Species are grouped by family with continent in parentheses. Metrics include IUCN conservation status (CS) with vulnerable (V), data deficient (DD), least concern (LC), and endangered (EN); mean alleles per locus ( $N_a$ ); allelic richness (AR); observed ( $H_o$ ) and expected ( $H_E$ ) heterozygosity. A “-“ indicates that genetic diversity metric was not reported.....59

**Table S2.** Primer sequences for 42 microsatellite loci optimized for the Caney Fork form of *Cambarus pristinus* and the Sequatchie form of *C. pristinus* including locus features: repeat motif, the number of alleles observed ( $N_a$ ), allele size range in base pairs (bp), and annealing temperature ( $T_a$ ). Data are summarized from the number of individuals indicated except for loci in bold which are summarized from our full data set of the Caney Fork form of *C. pristinus* with 169 individuals and an exploratory data set of the Sequatchie form with 21 individuals.....60

**Table S3.** Effective population size ( $N_e$ ) estimates generated herein and population census size ( $N_c$ ) estimates expressed as number of individuals/100 meters from Johansen et al. (2016). A “-“ indicates that the metric was not reported. Site ID corresponds to Table 1.....65

## LIST OF FIGURES

**Figure 1.** Known historical localities for *Cambarus pristinus*. The Caney Fork form is denoted by circles and the Sequatchie form is denoted by diamonds. A black “X” denotes an uncollected locality. Shapes filled with color denote sites where *C. pristinus* was detected and captured for use in our genetic assessments and unfilled shapes denote sites where *C. pristinus* was not detected. The white star denotes where the Caney Fork river becomes  $>4^{\text{th}}$  order. The Cumberland Plateau is represented in gray and the surrounding regions are white. The thick black line represents the Western Escarpment of the Cumberland Plateau. Letters by shapes correspond to Site ID defined in Table 1 .....34

**Figure 2.** The 95% statistical parsimony haplotype network of the 51 individuals of *C. pristinus* examined using the mitochondrial COI gene. Two networks were recovered, with each network corresponding to a single form of *C. pristinus*. Haplotypes differed by 1-4 mutations within each form. The number of mutations between haplotypes of each form exceeded the connection limit of 11 mutational steps. A circle represents a single haplotype and the “H” numbers designate different haplotypes. Size of the circles reflect the number of individuals sharing that haplotype. Colors represent individual sites and correspond to Figure 1. Open circles represent unsampled haplotypes and the line between circles denote a single mutation.....36

**Figure 3.** Maximum likelihood 50% majority rule consensus tree from 1000 bootstraps recovered for *Cambarus pristinus* from the mitochondrial COI gene. Sequence

divergences between species pairs are in parentheses. Colors identify localities and corresponds to those in Figure 1.....37

**Figure 4.** STRUCTURE analysis based on 19 microsatellite loci from 165 individuals across eight localities representing the Caney Fork form of *Cambarus pristinus*. a) Geographic distribution of the six population clusters inferred from the hierarchical STRUCTURE analyses (by colored polygons). Colors correspond to the inferred ancestral population depicted in the STRUCTURE plots (b-d). Letters by circles represent Site ID defined in Table 1. b) Initial STRUCTURE analysis with and without use of LOCPRIOR. c) Hierarchical STRUCUTRE plot using LOCPRIOR of Cluster A recovered in the initial STRUCTURE analysis. d) Hierarchical STRUCTURE plot using the LOCPRIOR for Cluster B recovered in the initial STRUCTURE analysis. b-d) Boxes designate localities and the color of each individual bar within a box denotes population ancestry of each individual. Letters below boxes and by circles correspond to site ID in Table 1. Delta K method is denoted by " $\Delta K$ "; the mean log-likelihood method is denoted by " $\mu L(K)$ ", and K represents the number of clusters inferred. An "\*" identifies the 6 distinct clusters recovered from all STRUCTURE runs.....38

**Figure 5.** Discriminant analysis of principle components (DAPC) based on 19 microsatellite loci and with individuals labeled by locality for the Caney Fork form of *Cambarus pristinus*. Ellipses denote the 6 genetic clusters recovered that correspond to the 6 clusters recovered in the STRUCTURE analyses (Figure 4). The top-right inset shows retained discriminant eigenvalues of the 5 principal components in relative

magnitude. The first two principal axes are shown, with the first on the vertical axis.....40

**Figure 6.** Isolation by distance analysis for the eight sites of *Cambarus pristinus* showing the relationship between geographic (log-transformed river and Euclidean kilometers) and genetic distances ( $F_{ST}$  and Jost's D) based on 19 microsatellite loci. For corresponding pairwise values refer to Table 3. Comparisons using log-transformed river distances are in black and those using log-transformed Euclidean distances are in gray (River-  $F_{ST}$ :  $R^2=0.0776$ ;  $p=0.1479$ , Jost's D:  $R^2=0.1266$ ;  $p=0.091$ ; Euclidean-  $F_{ST}$ :  $R^2=0.1570$ ;  $p=0.0542$ , Jost's D:  $R^2=0.1511$ ;  $p=0.0541$ ).....41

**Figure S1.** Criteria for selecting the number of retained principle components (PCs) and clusters for DAPC analysis. Red circles denote values retained in the DAPC analysis a) Graphic depicting recommended number of PCs to retain from the *optim.a* function in Adegenet. The a-score is the difference between the proportion of successful reassignment and values obtained using random groups retaining different numbers of PCs. b) Graph depicting number of retained PCs and explained cumulative variance. c) Graph depicting number of cluster and corresponding Bayesian information criterion (BIC) score. d) Proportion of individuals of a sampling site assigned to a genetic cluster denoted by DAPC (represented as the columns of boxes). The size of the squares represents the proportion of the individuals with the scale in the bottom left. Cluster letters correspond to Figure 5. e) Bar plot depicting membership probability of each individual to a cluster. Each box denotes a preassigned grouping by Site ID and color

corresponds to cluster color used in Figure 4b-d. Each bar represents an individual and membership probability is determined by the 5 retained discriminant eigenvalues expressed as a percent.....66

**Figure S2.** Maximum likelihood tree recovered for *Cambarus pristinus* from the recovered mitochondrial COI gene showing within clade phylogenetic relationships. Support values >50% bootstrap support are shown. Site ID corresponds to Table 1 and numbers in parentheses represent number of individuals from that site recovered at that node. Numbers in brackets show within clade sequence divergence.....68

**Figure S3.** Geographic distribution of the six distinct populations of the Caney Fork form inferred from the STRUCTURE and DAPC analyses in relation to specific anthropogenic disturbances (strip mining and silviculture) in the region. Colored polygons represent a distinct population cluster and the black circle inside denotes a collected locality. The size of the black circle denotes the allelic richness recovered for each site and corresponds to the scale in the bottom left.....69

## **Introduction**

Crayfishes are a diverse group of benthic, freshwater crustaceans with over 600 recognized species and new species described annually (Crandall and Buhay 2008). Approximately two-thirds of this diversity is found in the southeastern United States in the family Cambaridae and many species in this family exhibit high degrees of endemism (Lodge et al. 2000; Taylor et al. 2007; Dyer et al. 2013; Crandall and De Grave 2017; Dyer and Brewer 2018; Loughman and Williams 2018). Crayfishes, like many invertebrates, are a vital part of aquatic community structure. They act as ecosystem engineers by burrowing into substrates and creating microhabitats for other invertebrates (Creed and Reed 2004; Pintor and Soluk 2006; Glon and Thoma 2017), alter trophic interactions directly and indirectly via feeding links at multiple trophic levels (Creed 1994; Lodge et al. 1994; Dorn and Wojdak 2004), and are a vital part of many fish diets (Whitledge and Rabeni 1997). Despite their known biological importance and biodiversity there is a paucity of information on the distribution, life history, and genetic diversity of many species (Taylor and Schuster 2004; Moore et al. 2013; Figiel 2016), which impairs development of effective conservation management strategies for the estimated 48% of North American crayfishes considered imperiled (Taylor et al. 2007).

Overexploitation, habitat fragmentation, pollution, and invasive species are the main drivers of crayfish imperilment (Taylor et al. 2007). Several activities that contribute to habitat fragmentation and alteration in North America crayfishes include urban development, dams/water management, and logging/silviculture practices (Richman et al. 2015). Many crayfishes also have small geographic ranges and small



populations, which makes them more vulnerable to such human-mediated threats (Gaston 1994; Angermeier 1995).

In general, population connectivity via gene flow is disrupted by human-mediated fragmentation but is also influenced by several other factors, including dispersal ability of an organism, natural physical barriers in the environment and, geographic distance between populations (Koizumi et al. 2012; Phillipsen et al. 2015; Davis et al. 2018; Serrao et al. 2018; Tonkin et al. 2018). Compared to other aquatic groups such as fishes, few studies have examined population structure or the overall genetic diversity in crayfish species to explicitly test how such factors may impact population connectivity; the majority of those that have were conducted in Europe and Australia/Oceania (Figiel 2016). Studies that have focused on North American crayfishes have been limited to understanding genetic diversity and population connectivity of invasive populations, phylogeography of widely distributed species, or testing phylogenetic relationships within Cambaridae (Barbaresi et al. 2003; Fetzner and Crandall 2003; Finlay et al. 2006; Yue et al. 2010a). Given the high diversity and level of imperilment of North American crayfishes (Taylor et al. 2007), studies that examine how specific factors may influence population connectivity, especially for species that have small native ranges, are needed. Such studies can inform management of imperiled species by providing a basis for estimating the migration or dispersal potential of species, identifying historical processes that have shaped species distributions, predicting how species may respond to disturbance, and testing hypotheses about the relationships between species traits and the degree of population connectivity observed (Lowe and Allendorf 2010; Koizumi et al. 2012; Wang 2014; Wang and Bradburd 2014; Lowe et al. 2017). For example, it is often

hypothesized that species with low dispersal ability would exhibit high levels of population structure or reduced gene flow, as in the Murray Crayfish, *Euastacus armatus* of Australia (Whiterod et al. 2016). However, Finlay, et al. (2006) found a high degree of connectivity between surface and cave populations of the Cavespring Crayfish, *Cambarus tenebrosus*, despite a presumed limited dispersal ability in the species.

In riverine environments, dispersal is further limited due to the linear or dendritic pattern of these environments that restrict movement opportunities to narrow aquatic habitat corridors (Fagan 2002; Tonkin et al. 2018). Due to this, habitat degradation or artificial fragmentation is more likely to inhibit gene flow, resulting in population isolation, genetic drift, and reduced population persistence (Ward et al. 1994; Alp et al. 2012; Pilger et al. 2015, 2017). Given the largely linear nature of riverine systems, it is also expected that populations of species will typically display a pattern of increasing genetic isolation with increasing geographic distance among populations known as isolation-by-distance (IBD; Fetzner et al. 2003; Sexton et al. 2013).

Crayfishes, however, may be less constrained by their aquatic environments than other taxa and thus, may not show expected patterns of population isolation. For example, some species engage in limited overland dispersal (Furse et al. 2004; Oliveira et al. 2015; Lipták et al. 2016; Herrmann et al. 2018; Thomas et al. 2018) and dispersal ability is highly variable among crayfishes. Some species exhibit high degrees of dispersal moving across several aquatic systems (Barbaresi et al. 2004; Li et al. 2012; Loughman et al. 2013) and others show limited dispersal with individuals moving  $\leq 1$  km (Robinson et al. 2000; Bubb et al. 2006; Hurry et al. 2015; Whiterod et al. 2016). Dispersal potential may also be tied to fecundity and body size. For instance, in darters, species with higher

fecundity exhibit a higher dispersal potential (Turner et al. 1996). Fecundity is closely tied to body size in crayfishes as the number of eggs fertilized is determined by the amount of available surface area on the underside of the mother's abdomen where the eggs are attached (Distefano et al. 2013). Several barriers limiting crayfish dispersal have been identified including; waterfalls, dams, high gradient streams, and rapid changes in stream flow (Light 2003; Kerby et al. 2005; Bubb et al. 2006; Foster and Keller 2011) but the effects of these barriers on population connectivity vary by species.

To further examine factors that impact population connectivity and population persistence in crayfishes, we examined genetic diversity in the imperiled Pristine Crayfish, *Cambarus pristinus*. It is a headwater-stream adapted crayfish, endemic to approximately 374 km<sup>2</sup> on the Cumberland Plateau of Tennessee (Figure 1; Withers and McCoy 2005; Rohrback and Withers 2006; Johansen 2018) and occurs only in lower order stream reaches (< 4<sup>th</sup>) and tributaries of the Upper Caney Fork River, Bee Creek, Big Brush Creek, and Cane Creek systems of the Caney Fork and Sequatchie River systems (Rohrback and Withers 2006; Johansen et al. 2016). Within these tributaries, Williams et al. (2004) noted that occurrence was directly related to the presence of large flat rocks in shallow pools with low flow, leading to a patchy distribution within each tributary. It is unclear if pool habitat use is related to habitat specificity, congener competition (*C. parvovulus* and *C. sphenoides*), or some other factor. At several sites, *C. pristinus* occurs in low densities (less than 50 individuals per 100 m<sup>2</sup>; Johansen et al. 2016).

There are two distinct morphological forms of *C. pristinus*: the nominal form, which occurs only within the Caney Fork River system (referred to herein as the Caney

Fork form) and the Sequatchie form, whose range spans a drainage divide, found in the Sequatchie River system (Tennessee River drainage) and one tributary of the Caney Fork River system (Cumberland River drainage; Figure 1). The forms are allopatric and distinguishable based on differences in gonopod morphology (Rohrback and Withers 2006). Within the Caney Fork form's range, many tributaries of the Caney Fork River system (excluding those of the Upper Caney Fork and within the Bee Creek system that are on the Cumberland Plateau) join the Caney Fork mainstem after it flows off the West Escarpment of the Cumberland Plateau (Figure 1). The West Escarpment is characterized by stark drops in elevation where many streams have, through erosion, created confined, fast-flowing gulfs or ravines (Bouchard 1975). Such drastic changes in stream gradient may have created long-standing barriers or filters to gene flow among tributary populations of *C. pristinus* from these areas of the Caney Fork. Because *C. pristinus* is a headwater-stream specialist, it is likely that the larger order ( $\geq 4^{\text{th}}$  order) reaches of the Caney Fork River also may limit dispersal. Headwater-stream specialists are adapted to dispersing within small stream systems and may be unable to disperse across or within streams that have shifted in topography and hydrology as the river becomes larger (Hurry et al. 2015; Paz-Vinas et al. 2015; Schmidt and Schaefer 2018b). For example, two small-stream adapted fishes, including *Etheostoma basilare*, the Corrugated Darter, and *Etheostoma akatulo*, the Bluemask Darter, show evidence of reduced population connectivity or isolation among tributaries of the Caney Fork River, suggesting the mainstem Caney Fork River may serve as a barrier to their dispersal (Hollingsworth Jr. and Near 2009; Robinson et al. 2013). Similarly, the Sequatchie form is distributed across a drainage divide (Caney Fork River flows to the Cumberland River drainage and the

Sequatchie River flows to the Tennessee River drainage). For aquatic organisms in general, but especially those with presumed low dispersal potential, drainage divides often can be barriers to gene flow as seen in the White-clawed Crayfish, *Austropotamobius pallipes*, and the Barrens Topminnows, *Fundulus julisia* (Gouin et al. 2006; Hurt et al. 2017). It is important to evaluate how historical processes such as these have shaped gene flow, population structure, and distribution of a species as these processes can leave genetic signatures that are detected by genetic markers, such as microsatellites, used to assess more contemporary levels of population connectivity or isolation and thus, lead to incorrect conclusions about how contemporary processes have impacted the populations studied (Balkenhol et al. 2009; Zellmer and Knowles 2009; Davis et al. 2014; Epps and Keyghobadi 2015). Therefore, our first objective was to determine if drainage boundaries, physiographic breaks, changes in hydrology, or other geological features have created long-standing isolation among populations of *C. pristinus*. In addition, we assessed whether the two forms distinguished by morphological differences showed genetic divergence, representing distinct genetic lineages.

Recent surveys have reported loss of *C. pristinus* at several historical localities (Williams et al. 2004; Withers and McCoy 2005; Rohrback and Withers 2006; Johansen et al. 2016) often in conjunction with anthropogenic disturbances such as bridge construction. Population loss can contribute to increased isolation of extant populations, increasing the likelihood of further extirpation due to loss of genetic diversity through genetic drift (Frankham 1995a). Given these observations, our second objective was to identify contemporary patterns of genetic diversity and population structure in *C. pristinus*. Given the expectation of strong IBD patterns in riverine taxa, we expected

contemporary population structure would be influenced by the spatial arrangement of populations, but also by increased habitat alteration and anthropogenic disturbances leading to reduced population connectivity relative to observed historical connectivity. The small geographic range, small population size, and headwater-stream specialization, and potential habitat specificity suggests that *C. pristinus* has an increased susceptibility to isolation by habitat fragmentation. This study will provide a much-needed assessment of how this and other similar species may have responded to both historic geological processes or features of the environment and to contemporary anthropogenic activities that alter population connectivity.

## Methods

### *Sample collections*

*Cambarus pristinus* specimen collection occurred from October, 2017 to October, 2018 using standard seining and dip net methods (Parkyn 2016) at 13 of the 30 known historical localities in the Upper Caney Fork and Sequatchie River drainages (Table 1; Figure 1); other historical localities were not sampled due to lack of landowner permission or proximity to previously collected sites (Figure 1). In addition, four non-historical localities in the Sequatchie River drainage were sampled (Table 1). Genetic samples were collected by taking a specimen's chelae or the whole specimen in 95% ethanol for DNA preservation. Whole specimens were collected if chelae were reduced/absent or if total carapace length (TCL) was  $\leq 16$  mm to ensure enough tissue for future genetic analyses.

### *DNA extraction*

DNA was extracted from chela or abdominal tissue using a Qiagen DNEasy Blood and Tissue kit following standard manufacturer protocols. Because exoskeleton debris impaired our extracted DNA and resulted in PCR failure, we removed debris from all tissues under a microscope prior to extraction. Two elutions were used to generate a 180  $\mu\text{L}$  and 80  $\mu\text{L}$  volume of extracted DNA. DNA was quantified on a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific). Samples over 100ng/ $\mu\text{L}$  were diluted to a concentration of 30 ng/ $\mu\text{L}$  for microsatellite locus amplification. Any concentrations under 100 ng/ $\mu\text{L}$  were used undiluted in microsatellite locus amplification. No dilutions were created for mitochondrial gene sequencing.

### *Mitochondrial DNA data collection and analyses*

Five individuals per locality, except for Long Fork (**LF**) where only a single individual was collected and used, were amplified using Polymerase Chain Reaction (PCR) and sequenced for the mitochondrial DNA (mtDNA) Cytochrome Oxidase Subunit I (COI) gene using previously published primers (Folmer et al. 1994). PCR reactions used a 25.0  $\mu\text{L}$  total reaction volume with 2.00  $\mu\text{L}$  of individual DNA, 0.75  $\mu\text{L}$  of 25mM  $\text{MgCl}_2$  (New England Biolabs), 2.50  $\mu\text{L}$  of Standard 10X buffer (New England Biolabs), 0.50  $\mu\text{L}$  of 10mM dNTP (New England Biolabs), 1.00  $\mu\text{L}$  10pM primer LCO1490, 1.00  $\mu\text{L}$  10pM primer HCO2198, 0.25  $\mu\text{L}$  5000 U/mL Taq Polymerase (New England Biolabs), and 17.00  $\mu\text{L}$  of PCR water. Thermocycler conditions included 35 cycles of the following: 30s at 94°C, 30s at 50°C, and 90s at 72°C, which occurred after an initial denaturization step of 3 minutes at 94°C. These steps were followed by a final

extension of 5 minutes at 72°C. Sanger sequencing was performed by Yale's DNA Analysis Facility and resulting sequences were aligned and edited using CodonCode Aligner v8.0.1 (Soft Genetics). To minimize the potential for pseudogenes, methods recommended by Buhay (2009) and Song et al. (2008) were followed.

To examine phylogeographic relationships among the resulting COI haplotypes, haplotype networks were constructed in TCS v1.2.1 using a statistical parsimony analysis (Templeton et al. 1992) with a 95% connection limit. Final haplotype network graphics were created in TCSBeautifier (Murias dos Santo et al. 2016). Additionally, a phylogenetic tree was constructed in MEGA v 7.0.26 (Kumar et al. 2016) using a maximum likelihood analysis with 1000 bootstraps. The Hasegawa-Kishino-Yano with gamma distribution (HKY+G) substitution model (Hasegawa et al. 1985) was used as the parameter model of sequence evolution. This was determined by the Akaike information criterion in MEGA's model selection test consisting of 24 possible substitution models (Nei and Kumar 2000). The final graphic was created in FigTree v 1.4.4 (Rambaut 2012).

#### *Microsatellite genotype data collection*

Eighty-two species-specific primers were examined for successful amplification and variation in the Caney Fork Form of *C. pristinus*. Forty-two loci successfully amplified but only 19 were polymorphic and scoreable (Table S2). These 19 loci were placed into five panels, consisting of four loci with each locus assigned a unique M13 dye tag for the entirety of the study. PCR reactions consisted of 10.00 µL total volume including 1.00 µL of 10X concentrate standard Taq reaction buffer - Mg free (New England Biolabs), 1.20 µL 25mM MgCl<sub>2</sub> (New England Biolabs), 0.20 µL 10 mM dNTPs (New England Biolabs), 0.25 µL 10pM forward primer, 0.50 µL 10pM reverse primer, 0.10 µL 5000 U/mL Taq



Polymerase, 0.10  $\mu$ L M13-labeled dye (Applied Biosystems, Inc.), and 1.00  $\mu$ L DNA. Thermocycler conditions consisted of 35 cycles of the following: 30s at 94°C, 30s at the primer specific annealing temperature, and 90s at 72°C, which occurred after an initial denaturization step of 1 minute at 94°C. These steps were followed by a final extension of 5 minutes at 72°C and held at 12°C until the PCR product could be retrieved.

The PCR products were multiplexed by panel and genotyped by the University of Florida Interdisciplinary Center for Biotechnology Research Genotyping and Gene Expression Core using the Liz 600 size standard and an AB3730 sequencer. The resulting genotypes were automatically scored in GeneMarker v1.6 (SoftGenetics), followed by manual confirmation and edits as needed. Microsatellite data were not generated for the Sequatchie form due to small sample sizes.

#### *Microsatellite marker validation*

Potential scoring errors of microsatellite loci due to null alleles was evaluated using PopGenReport v3.0.4 (Adamack and Gruber 2014) in R; scoring errors due to large allele dropout and stutter were evaluated with MICRO-CHECKER v2.2.3 (Oosterhout et al. 2004). Null allele validation was conducted following methods of Waples and Manangwa et al. (2018, 2019). Linkage disequilibrium for all locus pairs and departures from Hardy-Weinberg equilibrium (HWE) per locus were tested with GENEPOP v1.1.2 (Rousset 2008) in R. Parameters for HWE and linkage disequilibrium were evaluated with the default Markov chain parameter settings of 10,000 dememorization steps, 1,000 batches, 10,000 iterations per batch and p-values were adjusted following Bonferroni corrections to reduce Type I errors (Rice 1989).

### *Spatial genetic structure*

Genetic differentiation was assessed using two pairwise metrics as recommended by Meirmans and Hedrick (2011). Pairwise  $F_{ST}$  (Weir and Cockerham 1984) values were generated in GENETIX v4.05 (Belkhir et al.) and tested for significance using 10,000 permutations of the data. Pairwise Jost's D (Jost 2008; Jost et al. 2018) values were generated in PopGenReport v3.0.4 in R and tested for significance in GenAlEx v6.503 (Peakall and Smouse 2006, 2012).

Pairwise  $F_{ST}$  and Jost's D values were used to test for a signature of isolation by distance (IBD) using two pairwise geographic distance matrixes measured in Google Earth. One matrix consisted of pairwise log-transformed riverine distances (km). The second matrix was based on pairwise log-transformed Euclidean distance (km) to account for potential underground water flow between localities due to the Cumberland Plateau's karst topography (Buhay and Crandall 2005; Finlay et al. 2006) or direct overland movements (Oliveira et al. 2015; Thomas et al. 2018). Significance was tested between geographic distance and genetic distance using a Mantel test with 10,000 randomizations of the data using IBD v1.52 (Bohonak 2002). Partial Mantel tests were used to assess if the physiographic break of the Cumberland Plateau escarpment or shift to a large mainstem river (>4<sup>th</sup> order; Figure 1) had any effect on genetic distance. Partial mantel tests were conducted using only the pairwise river distance measures.

STRUCTURE v2.3.4 (Pritchard et al. 2000) was used to further examine population sub-division. The initial run consisted of the following parameters: no *a priori* population assignment, a generalized admixture model, correlated allele frequencies, 5 iterations for each value of  $K$  ( $K=1-9$ ), and 10,000 burn-in Markov chain Monte Carlo

(MCMC) steps followed by 100,000 MCMC steps. The most likely number of population clusters was determined by using the mean log-likelihood ( $\ln[\Pr(X|K)]$ ) (Pritchard et al. 2000) and  $\Delta K$  (Evanno et al. 2005) methods in the program STRUCTURE HARVESTER web v0.6.94 (Earl and VonHoldt 2012).

Hierarchical analyses (Vähä et al. 2007; Janes et al. 2017) of initial clusters were run in STRUCTURE using the same parameters as the initial run. Subsequent analyses were run in STRUCTURE using sampling localities as priors in the LOCPRIOR model as recommended by Hubisz et al. (2009) to assess potential weak population structure. The program CLUMPAK v1.1 (Kopelman et al. 2015) was used to summarize and present the output of independent runs of each  $K$ .

Discriminant analysis of principle components (DAPC) can describe between-subpopulation variation while minimizing within-subpopulation variation noise. DAPC was conducted in Adegnet v2.1.1 in R (Jombart 2008). The optimal number of clusters  $K$  was determined using the k-means procedure of the function *find.cluster* and inferred from the Bayesian Information Criterion (BIC; Figure S1). Parameters were established using 10,000 iterations and 100 randomly chosen starting centroids to allow the algorithm to converge. Five Principal Components (PC) were retained as determined by the function *optim.a.score* with 10 simulations.

#### *Genetic diversity estimation*

Genetic diversity metrics for each distinct population cluster recovered from our structure analyses of the Caney Fork form were calculated in the R package PopGenReport v3.0.4. Metrics calculated for each cluster included mean number of alleles per locus ( $N_a$ ), percent of polymorphic loci, allelic richness (AR), private allelic

richness (PAR), and observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity. The inbreeding coefficient ( $F_{IS}$ ) was determined using GENETIX v4.05 (Belkhir et al. 2004) with 10,000 permutations to test for significance. All metrics calculated, except for PAR, were also estimated for the Caney Fork form overall.

Effective population size ( $N_e$ ) was estimated using the linkage disequilibrium method in NeESTIMATOR v2.01 (Waples and Do 2010; Do et al. 2014) to account for a single year dataset. Effective population size was estimated per cluster and for the Caney Fork form with the upper and lower bounds determined by 95% CI and alleles with a frequency of  $<0.02$  excluded to prevent upward bias of  $N_e$  (Waples 2006).

The program BOTTLENECK v1.2.02 (Piry et al. 1999) was used to test for recent population declines in each cluster and for the Caney Fork form. An excess of heterozygosity was tested using a two-phase model (TPM) with 0, 10, and 20% multistep mutation rates (Zachariah Peery et al. 2012) with 36% variation (Di Rienzo et al. 1994) via a Wilcoxon sign-rank test for significance.

## Results

### *Sample collections*

We obtained tissue samples from 10 of the 13 historical localities and one non-historical locality. A total of 226 individuals were collected. At least 20 individuals were collected per site except for Little Laurel Creek (*LLC*), Spring Creek (*SC*), Flatrock Branch (*FB*), and Long Fork (*LF*; Table 1). All localities were included in our mitochondrial analyses to assess phylogeographic relationships. Only localities containing the Caney Fork form were used for our microsatellite analyses as explained in our methods.

### *Mitochondrial DNA data collection and analyses*

We recovered 12 unique haplotypes from the 51 individuals of *C. pristinus* sequenced for the mitochondrial COI gene (Figure 2). Eight haplotypes were recovered for the 40 Caney Fork form individuals and four haplotypes were recovered for the 11 Sequatchie form individuals. Two unique haplotype networks were recovered from the statistical parsimony analysis with a maximum connection limit of 11 steps (95%). One network consisted of the Caney Fork form and the second was comprised of the Sequatchie form (Figure 2). One common haplotype was shared among all but one locality, *SC*, of the Caney Fork form (H1 in Figure 2). A common haplotype was shared among 2 localities of the Sequatchie Form (H11 in Figure 2). Haplotypes within the Caney Fork form network differed by one to four mutations and those within the Sequatchie form by one to five mutations; between networks haplotypes differed by 11 or more mutations (Figure 2).

Similarly, two distinct clades, each representing a form, were recovered with >95% bootstrap support from our maximum likelihood analysis (Figure 3). There was low divergence and no definable geographic structure within forms (Figure S1). Sequence divergence estimates from uncorrected p-distances was greatest between forms at 2.3% (Figure 3). Sequence divergence within the Caney Fork form was lowest at 0.12% and sequence divergence within the Sequatchie form was similarly low at 0.18% (Figure S2).

### *Microsatellite marker validation*

A total of 169 individuals were genotyped for the Caney Fork form for the 19 loci examined. Of these, 165 individuals were successfully genotyped at 16 or more loci and were retained for subsequent analyses (53/3135 missing genotypes; 1.8%). No evidence of scoring error, stutter, or allele dropout was found in any locus. Evidence of null alleles was found in 5 loci but at only 1–2 localities. All 19 loci were retained for subsequent analyses due to a lack of linkage disequilibrium across sites, no departures from HWE among loci, and the inconsistency of null allele among sites.

### *Spatial genetic structure*

Both the initial STRUCTURE analysis without *a priori* population assignments and the run using our eight Caney Fork form sites as priors in a LOCPRIOR model recovered  $K=3$  as the most likely number of clusters using the mean log-likelihood method (Figure 4b). The  $\Delta K$  method for the initial STRUCTURE analysis also indicated  $K=3$  but recovered  $K=2$  under our LOCPRIOR model. Due to the congruence of the initial analysis under both methods and the mean log-likelihood method under the LOCPRIOR model, a  $K=3$  including Cluster A, Cluster B, and Cluster SC (Figure 4b) was used for the subsequent hierarchical analyses. Little to no admixture was observed among these three clusters (Figure 4b).

Potential hidden structure was further examined using hierarchical STRUCTURE analyses for clusters recovered from our initial STRUCTURE output that included multiple sites. Cluster A, which included sites *CF*, *LLC*, *PoC*, *MC* and *WF* was examined using the LOCPRIOR model in which each site was used as a prior. The  $\Delta K$  method indicated  $K=2$  and the mean log-likelihood method indicated  $K=3$ , but both

recovered the West Fork (**WF**) as a distinct cluster. An additional STRUCTURE analysis excluding **WF** from Cluster A was conducted to determine if there was additional hidden structure in Cluster A; both the  $\Delta K$  and mean log-likelihood methods recovered  $K=2$ , corresponding to site Meadow Creek (**MC**) and the Upper Caney Fork clusters (Figure 4c). Cluster A was ultimately split into three clusters with sites **MC** and **WF** as distinct clusters and sites Caney Fork (**CF**), Pokepatch Creek (**PoC**), and Little Laurel Creek (**LLC**) within the third cluster (Upper Caney Fork cluster), with evidence of only low levels of admixture among these three clusters (Figure 4c). The hierarchical STRUCTURE analysis for Cluster B had congruence between the  $\Delta K$  and mean log-likelihood methods in that both recovered  $K=2$ . These two clusters represented sites **WFLCC** and **PuC** which had little to no admixture (Figure 4d). Ultimately, we recovered six distinct populations clusters denoted as: Upper Caney Fork Cluster, MC, WF, PuC, WFLCC, and SC clusters (Figure 4); most clusters were comprised of a single locality examined (Figure 4a).

Six distinct clusters, using 5 principle components (PCs), also were inferred from the discriminant analysis of principle components (DAPC; Figure S1a; b; c). These six clusters were similar to the six distinct population clusters recovered from our STRUCTURE analyses except that some individuals from the Upper Caney Cluster were assigned to either the MC or WF cluster in our DAPC analysis; this discrepancy is likely due to the limits of the membership assignment algorithm (Figure S1d; e).

Pairwise  $F_{ST}$  (range: 0.002–0.511) and Jost's D values (range: 0.002–0.442) were similar across several comparisons and the majority were significant after Bonferroni correction. Spring Creek (**SC**) consistently had the highest values under both metrics

indicating that **SC** is the most genetically differentiated site (Table 3). The few sites where pairwise comparisons were not significant were between sites found within the Upper Caney Cluster recovered from our STRUCTURE and DAPC analyses providing support for connectivity among the sites within the Upper Caney Fork Cluster and isolation of all other sites, resulting in the 6 distinct population clusters recovered.

We found no significant pattern of IBD using river distance ( $F_{ST}$ :  $R^2=0.0776$ ;  $p=0.1479$ , Jost's D:  $R^2=0.1266$ ;  $p=0.091$ ) and Euclidean distance ( $F_{ST}$ :  $R^2=0.1570$ ;  $p=0.0542$ , Jost's D:  $R^2=0.1511$ ;  $p=0.0541$ ). In all analyses, <16% of the total genetic variation was explained by geographic distance between sites (Figure 6). Our partial mantel tests showed a significant correlation between the West Escarpment of the Cumberland Plateau and our pairwise genetic metrics (Jost's D:  $R^2=0.7688$ ,  $p=0.0001$ ;  $F_{ST}$ :  $R^2=0.6281$ ,  $p=0.0001$ ) and a non-significant correlation between the larger order of the Caney Fork River and our pairwise genetic metrics (Jost's D:  $R^2=0.0568$ ,  $p=0.1462$ ;  $F_{ST}$ :  $R^2=0.0459$ ,  $p=0.1080$ ). However, our relationship between geographic and genetic distance was still non-significant after accounting for the effect of the West Escarpment of the Cumberland Plateau ( $F_{ST}$ :  $R^2=0.0776$ ;  $p=0.1582$ , Jost's D:  $R^2=0.0874$ ;  $p=0.1494$ ).

#### *Genetic diversity estimates*

A total of 115 alleles were amplified across all loci with an average of 6.1 alleles per locus across all clusters (Range: 2–12 alleles). Allelic diversity was low overall, but relatively uniform among clusters, with  $N_a$  ranging from 2.16–4.84 alleles and  $AR$  ranging from 1.89–4.03 alleles (Table 2). Observed ( $H_o$ ) and expected heterozygosity ( $H_e$ ) were similar across all clusters (range: 0.33–0.51  $H_o$ ; 0.38–0.54  $H_e$ ) except for **PuC** ( $H_o=0.16$ ;  $H_e=0.20$ ; Table 2), which had relatively lower heterozygosity. Private allelic



richness (PAR) was  $<0.10$  except for *WFLCC* and *SC*. Half of the clusters did not have  $F_{IS}$  values that differed significantly from zero but the other half had significant positive  $F_{IS}$  values indicating the presence of inbreeding for those clusters (Table 2). A signature of a bottleneck event was detected for *PuC* with a 20% two phase model (TPM). There was no evidence for deviations from HWE after Bonferroni correction within any cluster (Table 2).  $N_e$  varied among clusters (64.5–1347.5) but was low overall for the Caney Fork form with an estimate of 15.5 individuals (95% CI: 14.2–16.9).

## Discussion

Benthic, headwater-adapted species often have relatively low dispersal potential and exhibit hierarchical population structure due to natural instream conditions such as shifts in stream topography and hydrology (Hurry et al. 2015; Schmidt and Schaefer 2018b). Additionally, for such species, human-mediated habitat fragmentation often has a greater impact on the persistence of populations, and thus the species overall, due to their already restricted dispersal potential and limited dispersal pathways (Ward et al. 1994; Alp et al. 2012; Paz-Vinas et al. 2015).

For conservation management, it is important to disentangle population structure generated by geological features from structure created by anthropogenic disturbance. This is because structure related to anthropogenic habitat degradation or fragmentation can be confounded by or confused with signals from long-standing isolation from natural instream filters or barriers to gene flow (Zellmer and Knowles 2009; Davis et al. 2014; Epps and Keyghobadi 2015). Thus, we examined the potential for historic long-standing isolation within *C. pristinus* due to natural instream factors such as physiographic breaks and changes in river size, or past vicariance, and the natural degradation of genetic

similarity due to geographic distance known as isolation-by-distance (IBD). Given human-mediated habitat degradation and population extirpation observed in *C. pristinus*, we also evaluated contemporary population connectivity. We found that historical connectivity was maintained within the nominate form of *C. pristinus* (Caney Fork form) despite potential geological and hydrological barriers but found signatures of increased population isolation under contemporary processes. Contemporary signatures of increased isolation were not attributed to IBD. We also found evidence of divergence between the nominal Caney Fork form and the Sequatchie form, supporting the hypothesis that these represent distinct taxa.

#### *Phylogeography of Cambarus pristinus*

Previous studies have shown that large mainstem portions of rivers, changes in geology through physiographic breaks, and drainage boundaries serve as filters or barriers to dispersal of various benthic, aquatic taxa and that these barriers lead to long-standing isolation (Hughes et al. 1995; Fetzner and Crandall 2003; Nguyen et al. 2004; Hollingsworth Jr. and Near 2009; Kanno et al. 2011; Lamphere and Blum 2011; Hurry et al. 2015). Therefore, we expected to find signatures of long-standing isolation within *C. pristinus* attributed to reduced dispersal across a drainage divide, across the larger river portions of the mainstem Caney Fork River, and changes in stream gradient associated with the physiographic break of the Western Escarpment of the Cumberland Plateau. Although we found divergence between the two forms, suggesting their long-standing isolation, we found no evidence to suggest long-standing isolation within either form due to these presumed barriers. Each form had a common haplotype that was shared among the majority of sampled localities and haplotype divergence within each form was low;

together these data indicate that gene flow was likely maintained among populations at some level historically. This contrasts with the high degree of divergence exhibited by *E. basillare*, a Caney Fork River endemic. Like *C. pristinus*, *E. basillare* are benthic, and small-stream adapted with a presumed limited dispersal potential, particularly across large river habitats. Unlike *C. pristinus*, *E. basillare* displays a high degree of micro-endemism in which populations show evidence of isolation among tributary systems of the Caney Fork River. For example, each of several tributary systems to the Caney Fork (where *E. basillare* occurs) contains a different cryptic species of *E. basillare*; diverging between 8-2 mya. The authors attributed isolation of populations among tributaries to the strict breeding habitat requirements and reduced larvae dispersal potential of these species (Hollingsworth Jr. and Near 2009; Robinson et al. 2013; Fluker et al. 2014).

Crayfishes, however, may have less stringent habitat requirements for successful breeding. Ovigerous females of *Cambarus friauffi*, the Hairy Crayfish, opportunistically find brooding refuge (Black et al. 2015) and *Procambarus clarkii*, the Red Swamp Crayfish, readily disperses when brooding or caring for young (Oliveira et al. 2015). In addition, juvenile crayfish have been observed to disperse after becoming independent from the mother (Miller et al. 2014b; Glon et al. 2019). The relaxed breeding requirements and increased juvenile dispersal may explain why *C. pristinus* does not exhibit the same degree of isolation exhibited in darters of the Caney Fork River system. However, additional work is needed to test this explicitly. Sites **SC** of the Caney Fork form and **LF** of the Sequatchie Form each exhibited only a single haplotype unique to that site. For site **LF** this may suggest that the mainstem Big Brush Creek limits dispersal within the Sequatchie form although it is more likely a result of low sample size and

additional sampling would be needed to provide resolution. Site **SC** lies at the periphery of the geographic range of the Caney Fork form of *C. pristinus*. Populations found at the edge of a species' range typically are constrained by less favorable habitat that may reduce migration rates, increase isolation, and reduce genetic diversity (Eckert et al. 2008; Micheletti and Storfer 2017).

Our study is not alone in showing historic connectivity within crayfish on the Cumberland Plateau. The broadly distributed crayfishes, *Cambarus parvoculus* (Mountain Midget Crayfish), *C. jezerinaci* (Spiny Scale Crayfish), and *C. distans* (Boxclaw Crayfish), all exhibit similarly low haplotype diversity and divergence (Thoma and Fetzner, Jr 2008), however these species appear to be more generalist in their habitat requirements. Historic connectivity has been maintained in other headwater stream-adapted crayfish as well. *Euastacus bispinosus*, the Glenelg Spiny Crayfish, was shown to maintain historic connectivity across drainages in Australia (Miller et al. 2014a). Additionally, all headwater-stream specialists species may not exhibit the same dispersal limitations and additional factors such as density, drainage shape, and variation in large river distance separating headwater habitats should be considered (Schmidt and Schaefer 2018a)

The recovery of unique, divergent haplotypes that were sorted into two divergent clades, suggests long-standing isolation of the two morphologically distinct forms of *C. pristinus*. The distribution of these two forms, however, does not correspond to contemporary geographic breaks examined, given that the Sequatchie River form spans the Cumberland-Tennessee River drainage divide. This pattern of discord between divergence and contemporary geographic breaks have been observed in several other

aquatic species (Buhay et al. 2007; Berendzen et al. 2008; Wagner and Blanton 2017). This phylogeographic pattern may reflect one of several alternative scenarios. For example, ancestral populations within the Caney Fork River may have experienced past vicariance resulting in two forms in this system with subsequent invasion of the Sequatchie system through underground, overland, or headwater transfer. Alternatively, it may reflect isolation of a widespread ancestral species between the two drainages with subsequent transfer of the Sequatchie form into the Caney Fork River. Given the distance that separates the confluence of the Tennessee and Cumberland Rivers (in western Kentucky) and the absence of any records of the species in streams  $> 4^{\text{th}}$  order in size, it seems less plausible that the current distribution of the Sequatchie form is due to long-distance dispersal through these two drainages. Estimation of divergence times from taxa closely related to *C. pristinus* and also found on the Cumberland Plateau would suggest the two forms of *C. pristinus* diverged sometime during or after the Pleistocene (Crandall et al. 2015). It is likely divergence between the two forms is related to climate oscillations or shifts in stream gradient and drainage patterns, which have contributed to speciation of several lineages of freshwater taxa (Thornbury 1965; Near et al. 2001; Berendzen et al. 2003, 2008; Kozak et al. 2006). Estimation of divergence times for the two lineages would help link the observed patterns to specific geological events in the area and to evaluate and test for potential alternative scenarios that may have led to divergence of the two forms.

### *Contemporary genetic diversity*

Although not directly comparable due to variation in alleles examined, heterozygosity and allelic richness of the Caney Fork form of *C. pristinus* were similar to values observed for other imperiled crayfishes (Azuma et al. 2011; Gouin et al. 2011; Li et al. 2012; Gross et al. 2013; Miller et al. 2014a). Heterozygosity and allelic richness of several imperiled crayfishes, including *C. pristinus*, were relatively lower when compared with non-imperiled crayfishes such as *Procambarus clarkii*, the Red Swamp Crayfish (Li et al. 2012; Table S4). This suggests that *C. pristinus* has reduced adaptive potential when compared to non-imperiled crayfishes (Frankham 1995).

The Caney Fork form overall did not exhibit evidence of a bottleneck, however, site ***PuC*** (Puncheoncamp Creek), showed signatures of a bottleneck, suggesting a recent drastic decline in population size at this site. The Cumberland Plateau has experienced extensive sandstone and coal mining activity until the Surface Mining Control and Reclamation Act of 1977 (Schorr et al. 2006) and Puncheoncamp Creek historically had mining practices in its headwaters directly upstream of our sampling site (Figure S3; Moore 1985). Additionally, several abandoned mining sites have been repurposed for silvicultural practice. Withers and McCoy (2005) expressed that the majority of sub-watersheds within the range of *C. pristinus* drain large tracts of timber management land and several streams were not protected from the deleterious effects of sediment transport to the stream, which may result in loss of interstitial space. Crayfish density has been reported to decline in the presence of mining and silviculture activities due to the loss of interstitial space, increase of metal pollutants, and reduced canopy cover (Allert et al. 2012, 2013; Welsh and Loughman 2015). Allert (et al. 2012, 2013) found that the

greatest impacts to crayfish density occurred directly downstream of mining sites with reduced effects as distance from mines increased. Surface and deep anthracite coal-mining have been designated as major factors in the decline of the federally listed Guyandotte Crayfish, *Cambarus veteranus* (Loughman et al. 2016) and logging practices have been a concern for several other crayfishes, including the Piedmont Blue Burrowing Crayfish, *Cambarus harti* and the vulnerable Kiamichi Crayfish, *Faxonius saxatilis* (Jones and Bergey 2007; Helms et al. 2013).

Other *C. pristinus* sites may have experienced past declines in population size that were not detected by our bottleneck analysis, which relies on measures of heterozygote excess. This method requires a reduction of 50 – 80% of the effective population to detect bottlenecks two-thirds of the time, even when additional markers are included (Hoban et al. 2013). In addition, other factors such as pre-bottleneck genetic diversity, bottleneck persistence, when the bottleneck occurred, and population growth can influence or obscure genetic signals (Williamson-Natesan 2005; Zachariah Peery et al. 2012).

We found evidence of inbreeding at several sites (**WFLCC**, **PoC**, **SC**, and **PuC**; Table 2). This was not surprising for **PuC**, because it showed signatures of a bottleneck and inbreeding often occurs after a bottleneck event (Hedrick and Kalinowski 2000). However, inbreeding can occur in small population without a prior bottleneck event. All clusters except for **SC**, had inbreeding coefficients with 95% confidence intervals that spanned zero. Thus, we could not confidently conclude the occurrence of inbreeding at the population level (Table 2; Colegrave and Ruxton 2003). However, we recovered a significant inbreeding coefficient for the Caney Fork form ( $F_{IS} = 0.22620$ ) with a 95% confidence interval that did not span zero. This result is concerning because, in many

species, inbreeding often leads to reduced fitness and an increased extinction risk (Frankham 2005; Wright et al. 2008). However, crayfishes may have an inherent degree of inbreeding due to poor dispersal ability (Miller et al. 2014a) and dominant hierarchical social structure, where dominant male crayfishes have additional mating opportunities (Villanelli and Cherardi 1998; Tierney et al. 2000; Moore and Bergman 2005). However, several crayfishes of Cambaridae exhibit polyandry with 2-3 males siring a single clutch (Walker et al. 2002; Yue et al. 2010b; Kahrl et al. 2014) and polyandry is thought to offset inbreeding depression by increasing half-sib progeny and decreasing full-sib progeny (Tregenza and Wedell 2000; Cornell and Tregenza 2007). It is unlikely that the significant inbreeding coefficient for the Caney Fork form is only due to mating strategies since we observed low effective population size estimates, low allelic richness, and localized bottleneck events in the species as well.

In wild populations of animals, the breeding population ( $N_e$ ) is on average approximately 10% of the population census size (Frankham 1995b). Johansen et al. (2016) estimated a census size for several populations of the Caney Fork form of *C. pristinus*; estimates ranged from 45–168 individuals (Table S3). Our point estimates of effective population size ranged from 64.5–1347.5 but the upper bounds of our confidence intervals were infinite for many clusters reducing our confidence in the  $N_e$  estimates (Table 2). However, the Caney Fork form of *C. pristinus* had low point estimates of  $N_e$  for most clusters examined and for the species overall. Individual cluster estimates are likely impacted by low sample sizes, but the estimate for the species overall had narrow confidence intervals, around the point estimate of only 15.5 individuals. This low estimate likely reflects both a low number of individuals and low genetic diversity in



the breeding population. The 100/1000 rule recommends that a  $N_e$  of 100 be maintained to prevent inbreeding depression and an  $N_e$  of 1000 to maintain evolutionary potential (Frankham et al. 2014). Following this rule, our data suggests that the Caney Fork form of *C. pristinus* lacks a sufficient breeding population size or has insufficient genetic diversity in the breeding population to prevent inbreeding and has reduced evolutionary potential.

A common pattern observed in aquatic systems is an increase in genetic distance as geographic distance increases or IBD (Wright 1943). We did not find support for IBD when using river distance in the Caney Fork form of *C. pristinus* (Figure 6). It is possible that IBD is masked at the small spatial scale that is the extent of our study, especially if low levels of gene flow are maintained as suggested by the few non-significant pairwise  $F_{ST}$  and Jost's D values we observed (Table 3; Phillipsen et al. 2015; Menger et al. 2017). We continued to find a non-significant IBD even after accounting for the significant correlation between genetic distance and the West Escarpment of the Cumberland Plateau suggesting that something other than geographic distance, such as habitat degradation, is generating the observed population structure.

It is unlikely that *C. pristinus* disperses across land. We recovered a non-significant IBD relationship when using Euclidean distance and a lack of population connectivity across headwater systems from our STRUCTURE analysis. If aquatic organisms maintain geneflow via overland dispersal, then populations across headwater systems in close proximity would exhibit higher degrees of genetic similarity than population spaced far apart within a single headwater system. (Finn et al. 2007). We found significant pairwise  $F_{ST}$  values among our populations and a high degree of

dissimilar genetic diversity metrics between our two nearest sites (<2km) across land, *PuC* and *WFLCC*. In addition, we found greater similarity across several genetic diversity metric within the Upper Caney Fork headwater system than among the Upper Caney Fork and Bee Creek headwater systems. Even though overland dispersal can occur within crayfishes it likely differs among species. For instance, even in dry periods, several species of stream-dwelling crayfish choose to burrow into the hyporheic zone instead of dispersing across land to available water sources (Jones and Bergey 2007; DiStefano et al. 2009; Dyer and Brewer 2018).

The congruence between our pairwise  $F_{ST}$  values and the STRUCTURE and DAPC analyses provide support for six isolated populations of the Caney Fork form of *C. pristinus*: (1) Upper Caney, (2) MC, (3) WF, (4) PuC, (5) WFLCC, and (6) SC. This high degree of contemporary population isolation contrasts with our phylogeographic analysis that showed connectivity was maintained throughout the Caney Fork form's range historically. One explanation is that the retreating edge of the Western Escarpment has resulted in a loss of underground stream connectivity (Anthony and Granger 2007a) as indicated by cave abandonment along the tributaries of the Caney Fork River (Anthony and Granger 2004). Cave abandonment is estimated to have begun ~5Mya with subsequent abandonment occurring into the early Holocene (Anthony and Granger 2007b). Loss of underground stream connectivity has been documented in several cave-obligate fauna, including crayfishes along the escarpment lines of the Cumberland Plateau (Buhay and Crandall 2005; Finlay et al. 2006; Buhay et al. 2007; Niemiller et al. 2008; Niemiller and Zigler 2013).

However, given that our data suggests populations maintained some level of gene flow historically and cave abandonment occurred over a long geologic time span, it is unlikely loss of underground connectivity has contributed to contemporary population isolation. Instead, the high degree of contemporary population isolation observed most likely reflects human-mediated habitat alterations and subsequent population loss resulting in reduced extant population size and connectivity. Anthropogenic disturbance such as mining and silviculture practice have been reported throughout the range of *C. pristinus* with several anthropogenic sites near the six distinct population clusters we recovered (Figure S3). Johansen (2018) found that the likelihood of *C. pristinus* occupancy would decrease as evergreen forest cover or conductivity increased. Previous studies have shown that acid mine drainage from mining can increase conductivity and the bulk of silviculture practice on the Cumberland Plateau is from loblolly pine, an evergreen, plantations (McGrath et al. 2004; Schorr et al. 2013). Coal mining activity on the Cumberland Plateau has been reported to reduce salamander richness (Schorr et al. 2013; Muncy et al. 2014), alter macroinvertebrate assemblages (Gangloff et al. 2015), create local extirpation and species replacement in native fishes (Schorr et al. 2006), and reduce population density or create local extirpation in crayfishes (Allert et al. 2012, 2013; Welsh and Loughman 2015).

#### *Conservation implications for Cambarus pristinus*

The observed geographic separation, phylogenetic divergence, and morphological differences between the two forms of *C. pristinus* support recognition of the Sequatchie form as a distinct species and we recommend it be provisionally referred to as *Cambarus* aff. *pristinus*. Due to this, the geographic range of nominal *C. pristinus* (herein the Caney

Fork form) is restricted to the Upper Caney Fork and Bee Creek system of the Cumberland Plateau.

Although, a population genetics analysis is needed to assess genetic diversity and gene flow within *Cambarus* aff. *pristinus*, repeated failures to detect individuals at several localities is concerning and may indicate local extirpation. Several individuals we collected had symptoms of the lethal Porcelain Disease, Thelohaniasis. Porcelain Disease occurs in several decapod crustaceans and is caused by the microsporidian parasites of the genus *Thelohania* (El-Matbouli and Soliman 2006). Disease prevalence is thought to be linked to increased stress and habitat alteration (Imhoff et al. 2009). Loblolly pine plantations are found throughout the range of *Cambarus* aff. *pristinus*; these plantations can alter water chemistry and stream habitat and thus, the benthic community structure (McGrath et al. 2004). We propose that additional surveys be conducted to determine the extent of the geographic range of *Cambarus* aff. *pristinus* and that future monitoring to detect Porcelain Disease be incorporated. In addition, a thorough assessment of genetic diversity within *Cambarus* aff. *pristinus* is warranted to determine population persistence.

Low genetic diversity,  $N_e$  estimates, presence of inbreeding, and population structure exhibited within *Cambarus pristinus sensu stricto* (Caney Fork form) suggests that they have a reduced evolutionary potential and ability to weather future stochastic events (Frankham 1995a; Lowe and Allendorf 2010), warranting continued state protection. In addition, we observed a high degree of contemporary population isolation. This is in stark contrast to historic connectivity as discussed and implies human-mediated habitat alternations have likely impacted this species. However, additional work from a landscape genetics approach would provide a more robust test for effects of specific

landscape and environmental factors have contributed to the observed patterns of genetic diversity observed. *Cambarus pristinus sensu stricto* will also benefit from habitat and occupancy modeling, which would help to understand how abiotic factors have influenced population growth and persistence and outline favorable habitat conditions. Continued population genetic monitoring would establish trends in genetic diversity in the species and populations, which can be used to assess the actual impact of conservation actions. Additionally, because our analyses only provide a single snapshot of genetic diversity, adding time-series sampling would lead to more robust estimates of  $N_e$  (Gilbert and Whitlock 2015).

**Table 1.** Locality information and number of individuals of *Cambarus pristinus* captured. A “+” denotes new localities identified and sampled for *Cambarus pristinus*. Site ID corresponds to those used in all other tables and figures.

Form	Site ID	Stream	County	Date Collected	Latitude	Longitude	No. Captured
Caney Fork	<b>CF</b>	Caney Fork River	Cumberland	14-Oct-2017; 08-May-2018	35.99108	-85.18052	24
Caney Fork	<b>WF</b>	West Fork	Cumberland	14-Oct-2017; 08-May-2018	35.92435	-85.21365	29
Caney Fork	<b>MC</b>	Meadow Creek	Cumberland	14-Oct-2017; 08-Apr-2018	35.88828	-85.17790	27
Caney Fork	<b>WFLCC</b>	West Fork Little Cane Creek	Cumberland	14-Oct-2017; 08-May-2018	35.79953	-85.20750	23
Caney Fork	<b>PoC</b>	Pokepatch Creek	Cumberland	08-Apr-2018 ;08-May-2018	35.94527	-85.18546	30
Caney Fork	<b>LLC</b>	Little Laurel Creek	Cumberland	22-Jun-2018	35.84049	-85.12305	11
Caney Fork	<b>SC</b>	Spring Creek	Van Buren	23-Jun-2018	35.77052	-85.28789	19
Caney Fork	<b>PuC</b>	Puncheoncamp Creek	Cumberland	23-Jun-2018	35.83650	-85.22590	31
Caney Fork		Henderson Branch	Cumberland	23-Jun-2018	35.82225	-85.16087	0
Sequatchie	<b>CC</b>	Camp Creek	Van Buren	11-Jan-2018	35.65099	-85.33764	21
Sequatchie	<b>FB</b>	Flatrock Branch +	Sequatchie	09-Oct-2018	35.49406	-85.40233	12
Sequatchie	<b>LF</b>	Long Fork 2	Sequatchie	09-Oct-2018	35.50113	-85.40877	1
Sequatchie		Big Brush Creek +	Sequatchie/Bledsoe	08-Sep-2018; 08-Oct-2018	35.50017	-85.40623	0
Sequatchie		Laurel Fork +	Sequatchie	24-Jun-2018	35.52947	-85.43130	0
Sequatchie		Unnamed Tributary to Bird Fork +	Sequatchie	08-Sep-2018; 08-Oct-2018	35.48935	-85.43797	0
Sequatchie		Glady Fork	Sequatchie	09-Sep-2018; 09-Oct-2018	35.52543	-85.46857	0
Sequatchie		Long Fork 1	Sequatchie	09-Sep-2018; 09-Oct-2018	35.50195	-85.45370	0

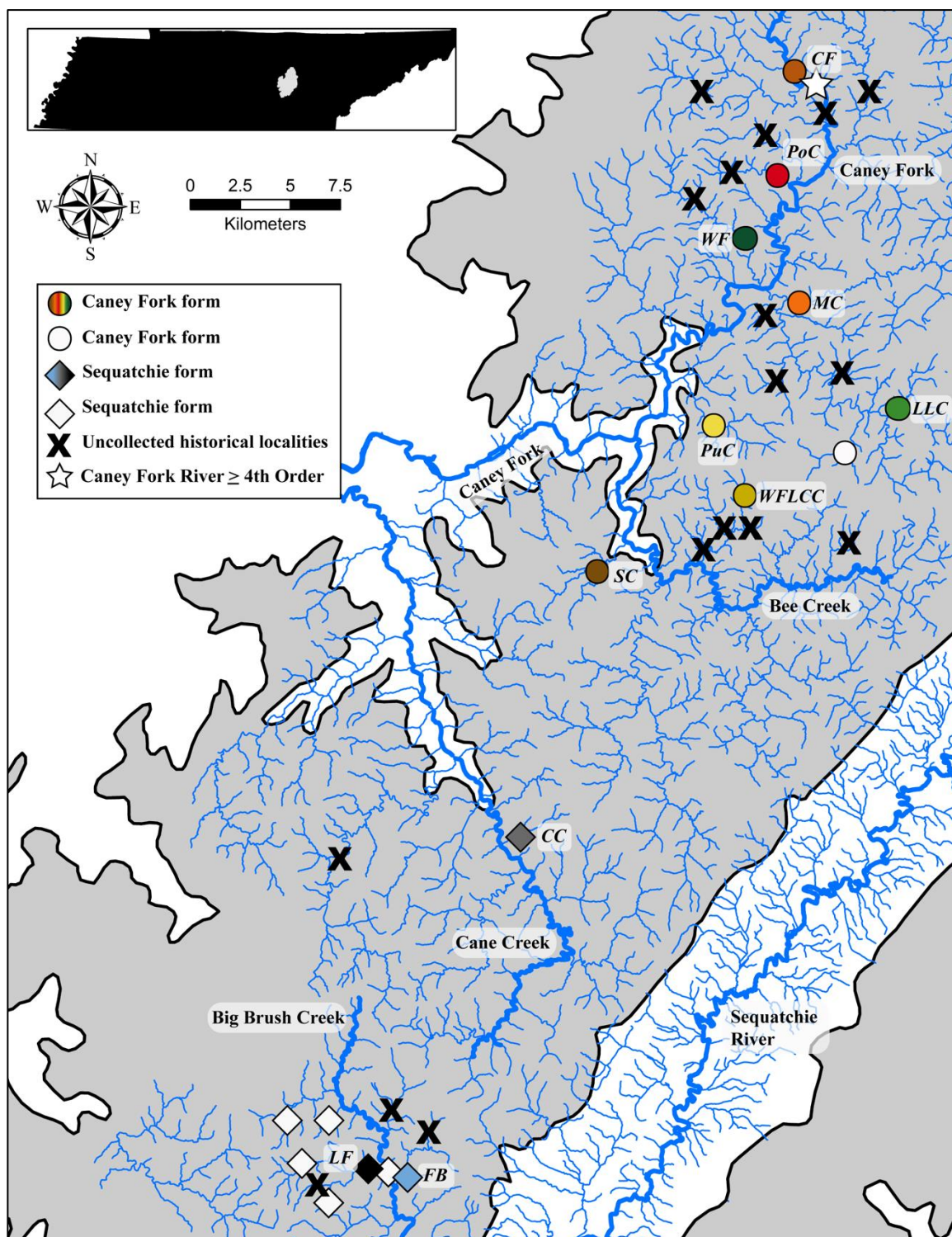
**Table 2.** Genetic diversity measures based on 19 microsatellite loci for 6 distinct population clusters examined for the Caney Fork form of *Cambarus pristinus*: mean alleles per locus ( $N_a$ ), allelic richness (AR), private allelic richness (PAR), observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity, inbreeding coefficient ( $F_{IS}$ ), p-values for Bottleneck, p-values for Hardy-Weinberg equilibrium (HWE), and percentage of polymorphic loci. Values in bold were significant ( $p < 0.05$  for  $F_{IS}$  and Bottleneck, and  $p < 0.003$  for HWE after Bonferroni correction); n is the number examined per distinct population cluster.

Cluster ID	n	$N_a$	$N_e$ (95% CI)	AR	PAR	$H_o$	$H_e$	$F_{IS}$ (95% CI)	Bottleneck	HWE	% Polymorphic loci
<i>Upper Caney</i>	58	3.95	$\infty$ (827– $\infty$ )	3.14	0.08	0.46	0.48	0.03237 (-0.025–0.074)	0.369	0.011	100%
<i>WF</i>	25	3.21	1046 (58.3– $\infty$ )	2.79	0.05	0.45	0.44	0.00318 (-0.100–0.059)	0.340	0.671	100%
<i>MC</i>	25	3.05	785.3 (53.1– $\infty$ )	2.56	0.00	0.44	0.43	-0.01705 (-0.124–0.045)	0.325	0.705	100%
<i>WFLCC</i>	22	4.84	136.6 (52.7– $\infty$ )	4.03	0.22	0.51	0.54	<b>0.08564</b> <b>(-0.007–0.128)</b>	0.384	0.007	100%
<i>SC</i>	19	2.16	1347.5 (19.3– $\infty$ )	1.89	0.12	0.16	0.20	<b>0.22550</b> <b>(0.069–0.300)</b>	0.892	0.009	68.42%
<i>PuC</i>	16	2.58	$\infty$ (26.9– $\infty$ )	2.34	0.06	0.33	0.38	<b>0.14904</b> <b>(-0.008–0.223)</b>	<b>0.039</b>	0.155	73.68%
<b>Species Overall</b>	165	6.10	15.5 (14.2–16.9)	2.66	-	0.41	0.42	<b>0.22620</b> <b>(0.179–0.265)</b>	0.964	0.197	100%

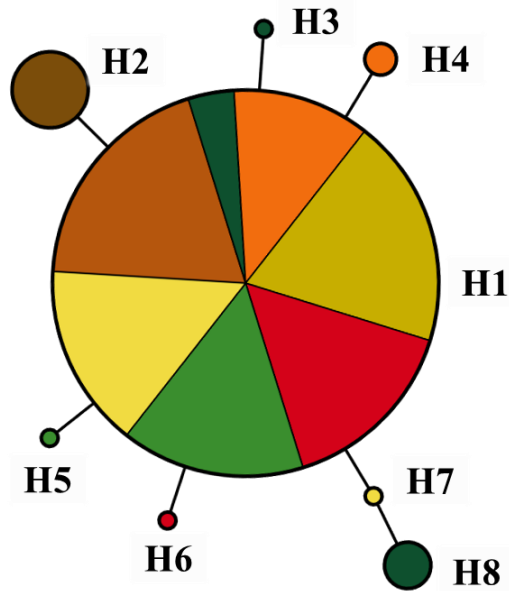
**Table 3.** Pairwise measures of genetic and geographic distances used for the IBD analysis of the Caney Fork form of *Cambarus pristinus* based on 19 microsatellite loci (Figure 4). Site ID corresponds to those used in Figure 1 and defined in Table 1. a) Pairwise  $F_{ST}$  values (above the diagonal) and pairwise Jost's D values (below the diagonal). Values in bold are significant at  $p < 0.05$  and values with an asterisk (\*) are significant following Bonferroni correction ( $p < 0.0018$ ). b) Pairwise river distance (above the diagonal) and pairwise Euclidean distance (below the diagonal) in kilometers.

a)	<b>Site ID</b>	<b>CF</b>	<b>WF</b>	<b>MC</b>	<b>WFLCC</b>	<b>PoC</b>	<b>LLC</b>	<b>SC</b>	<b>PuC</b>
	<b>CF</b>	-	<b>0.042*</b>	<b>0.062*</b>	<b>0.145*</b>	0.010	0.002	<b>0.412*</b>	<b>0.257*</b>
	<b>WF</b>	<b>0.038*</b>	-	<b>0.097*</b>	<b>0.171*</b>	<b>0.065*</b>	<b>0.046</b>	<b>0.420*</b>	<b>0.313*</b>
	<b>MC</b>	<b>0.054*</b>	<b>0.086*</b>	-	<b>0.181*</b>	<b>0.052*</b>	<b>0.034</b>	<b>0.426*</b>	<b>0.330*</b>
	<b>WFLCC</b>	<b>0.180*</b>	<b>0.209*</b>	<b>0.214*</b>	-	<b>0.143*</b>	<b>0.131*</b>	<b>0.333*</b>	<b>0.229*</b>
	<b>PoC</b>	0.010	<b>0.061*</b>	<b>0.046*</b>	<b>0.180*</b>	-	0.009	<b>0.441*</b>	<b>0.258*</b>
	<b>LLC</b>	0.002	<b>0.042*</b>	<b>0.029</b>	<b>0.170*</b>	0.010	-	<b>0.453*</b>	<b>0.283*</b>
	<b>SC</b>	<b>0.375*</b>	<b>0.379*</b>	<b>0.372*</b>	<b>0.322*</b>	<b>0.442*</b>	<b>0.394*</b>	-	<b>0.511*</b>
	<b>PuC</b>	<b>0.266*</b>	<b>0.336*</b>	<b>0.347*</b>	<b>0.279*</b>	<b>0.275*</b>	<b>0.303*</b>	<b>0.433*</b>	-
b)	<b>Site ID</b>	<b>CF</b>	<b>WF</b>	<b>MC</b>	<b>WFLCC</b>	<b>PoC</b>	<b>LLC</b>	<b>SC</b>	<b>PuC</b>
	<b>CF</b>	-	23.82	21.28	55.14	10.09	35.01	44.81	35.91
	<b>WF</b>	7.94	-	5.94	40.86	14.85	19.75	30.53	21.63
	<b>MC</b>	11.30	5.16	-	37.26	12.31	13.73	26.93	18.03
	<b>WFLCC</b>	21.30	13.90	10.10	-	46.14	51.07	15.07	26.95
	<b>PoC</b>	5.00	3.38	6.48	16.30	-	24.38	35.81	26.94
	<b>LLC</b>	17.40	12.40	7.16	8.88	12.80	-	40.74	31.84
	<b>SC</b>	26.20	18.40	16.40	7.93	21.50	16.70	-	16.62
	<b>PuC</b>	17.50	9.84	7.27	4.36	12.60	9.25	9.30	-

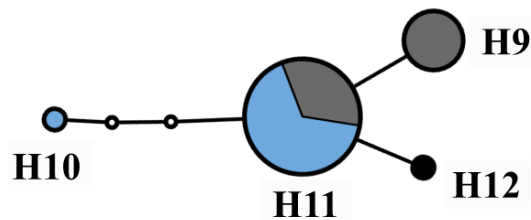




**Figure 1.** Known historical localities for *Cambarus pristinus*. The Caney Fork form is denoted by circles and the Sequatchie form is denoted by diamonds. A black “X” denotes an uncollected locality. Shapes filled with color denote sites where *C. pristinus* was detected and captured for use in our genetic assessments and unfilled shapes denote sites where *C. pristinus* was not detected. The white star denotes where the Caney Fork river becomes  $>4^{\text{th}}$  order. The Cumberland Plateau is represented in gray and the surrounding regions are white. The thick black line represents the Western Escarpment of the Cumberland Plateau. Letters by shapes correspond to Site ID defined in Table 1.

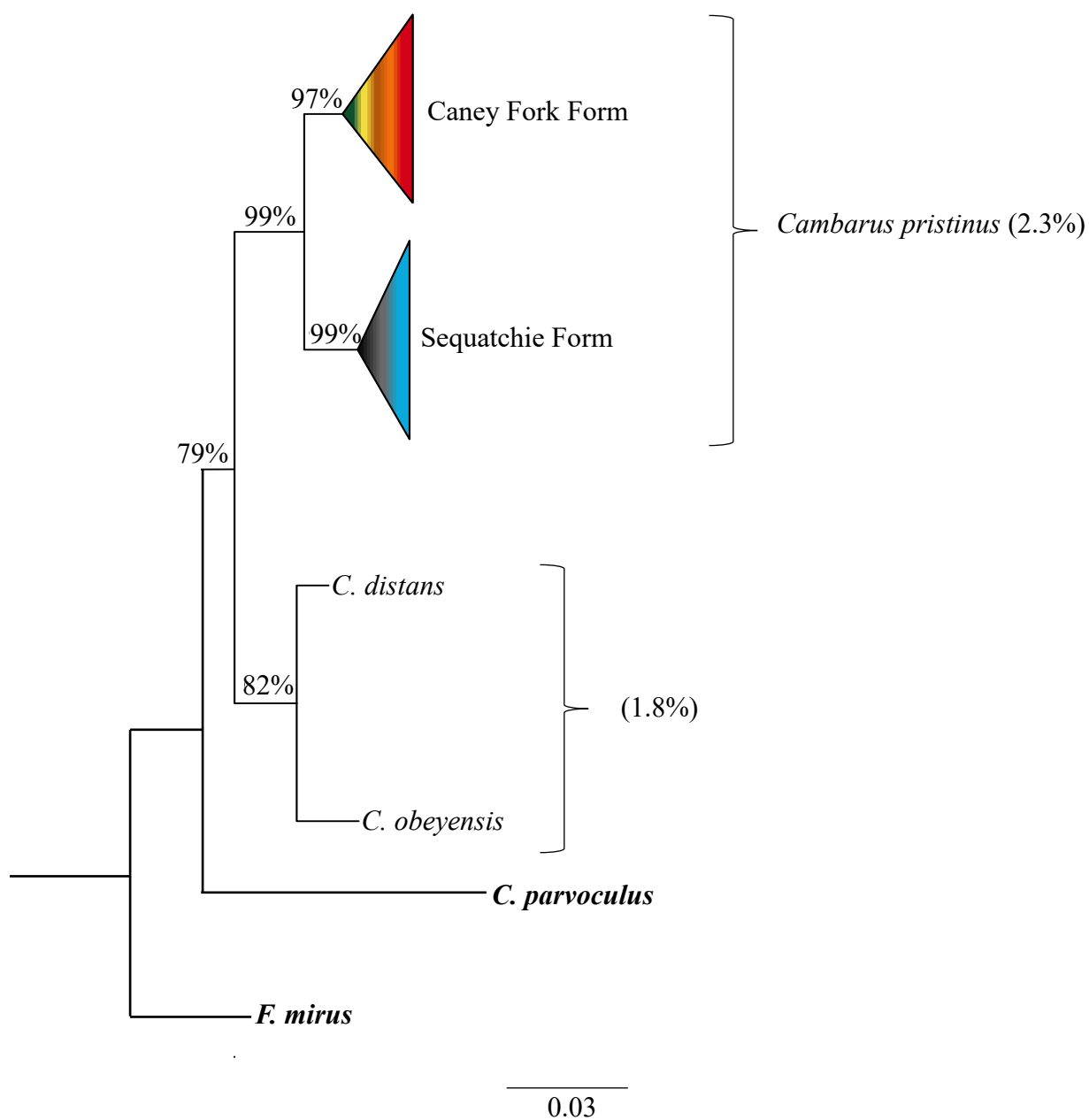


*C. pristinus* – Caney Fork form

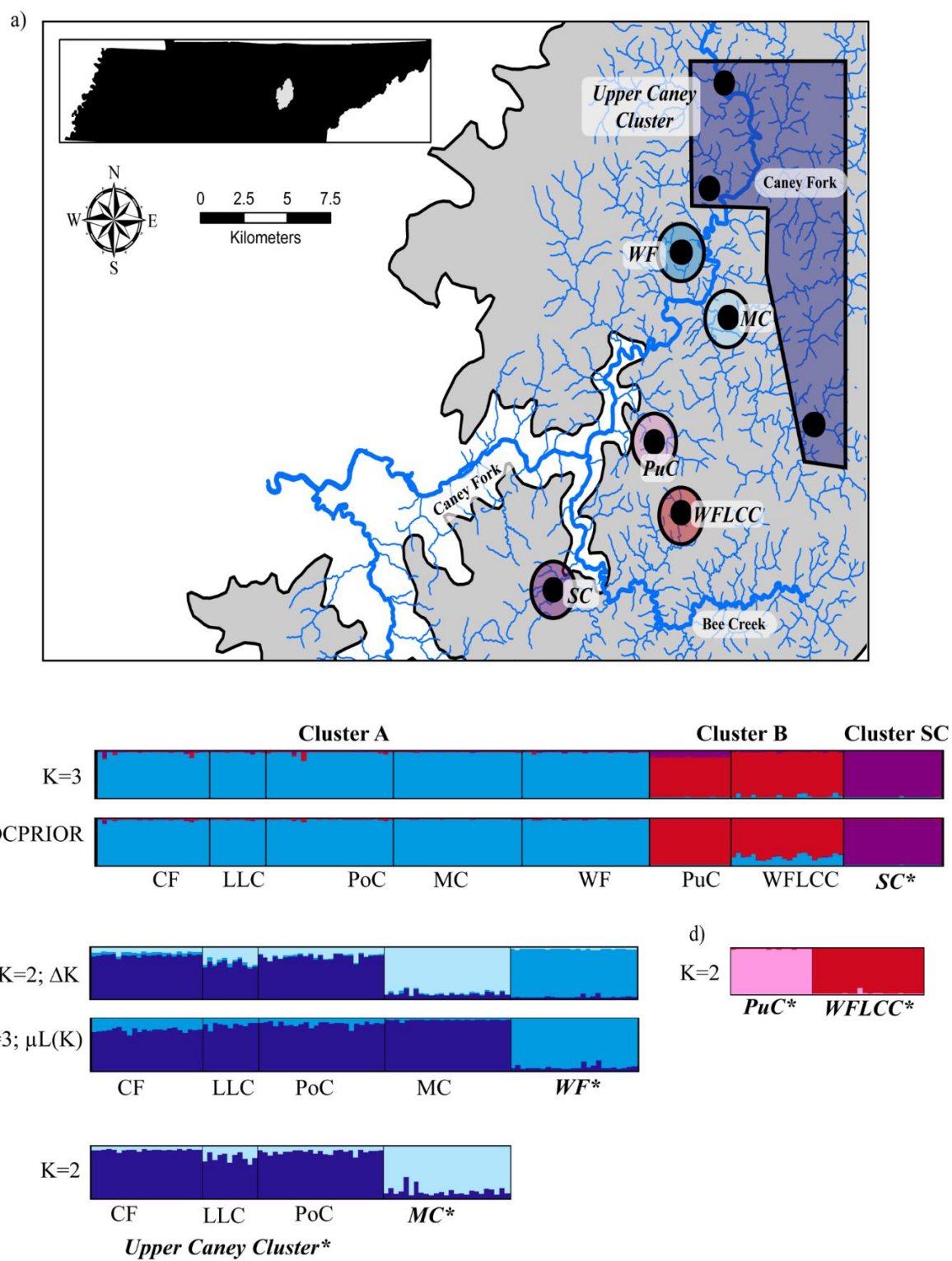


*C. pristinus* – Sequatchie form

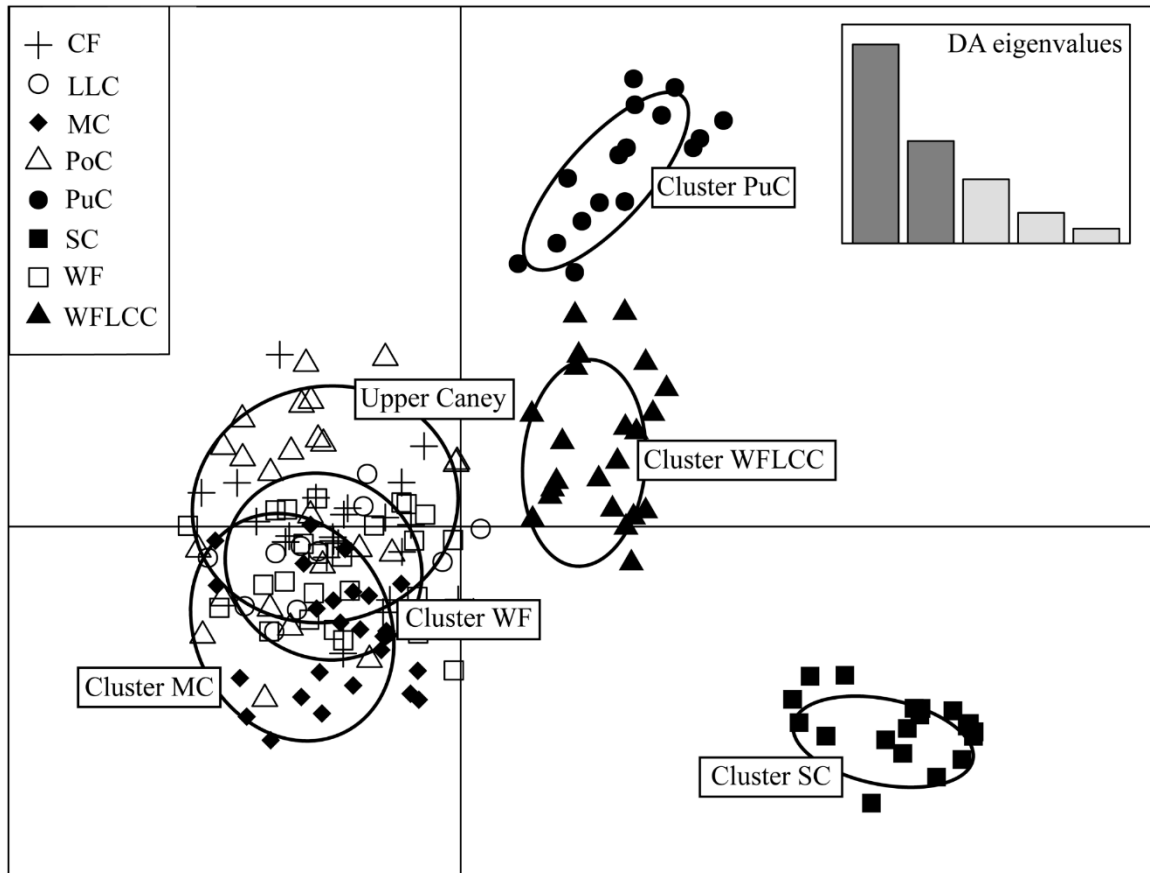
**Figure 2.** The 95% statistical parsimony haplotype network of the 51 individuals of *C. pristinus* examined using the mitochondrial COI gene. Two networks were recovered, with each network corresponding to a single form of *C. pristinus*. Haplotypes differed by 1-4 mutations within each form. The number of mutations between haplotypes of each form exceeded the connection limit of 11 mutational steps. A circle represents a single haplotype and the “H” numbers designate different haplotypes. Size of the circles reflect the number of individuals sharing that haplotype. Colors represent individual sites and correspond to Figure 1. Open circles represent unsampled haplotypes and the line between circles denote a single mutation.



**Figure 3.** Maximum likelihood 50% majority rule consensus tree from 1000 bootstraps recovered for *Cambarus pristinus* from the mitochondrial COI gene. Sequence divergences between species pairs are in parentheses. Colors identify localities and corresponds to those in Figure 1.

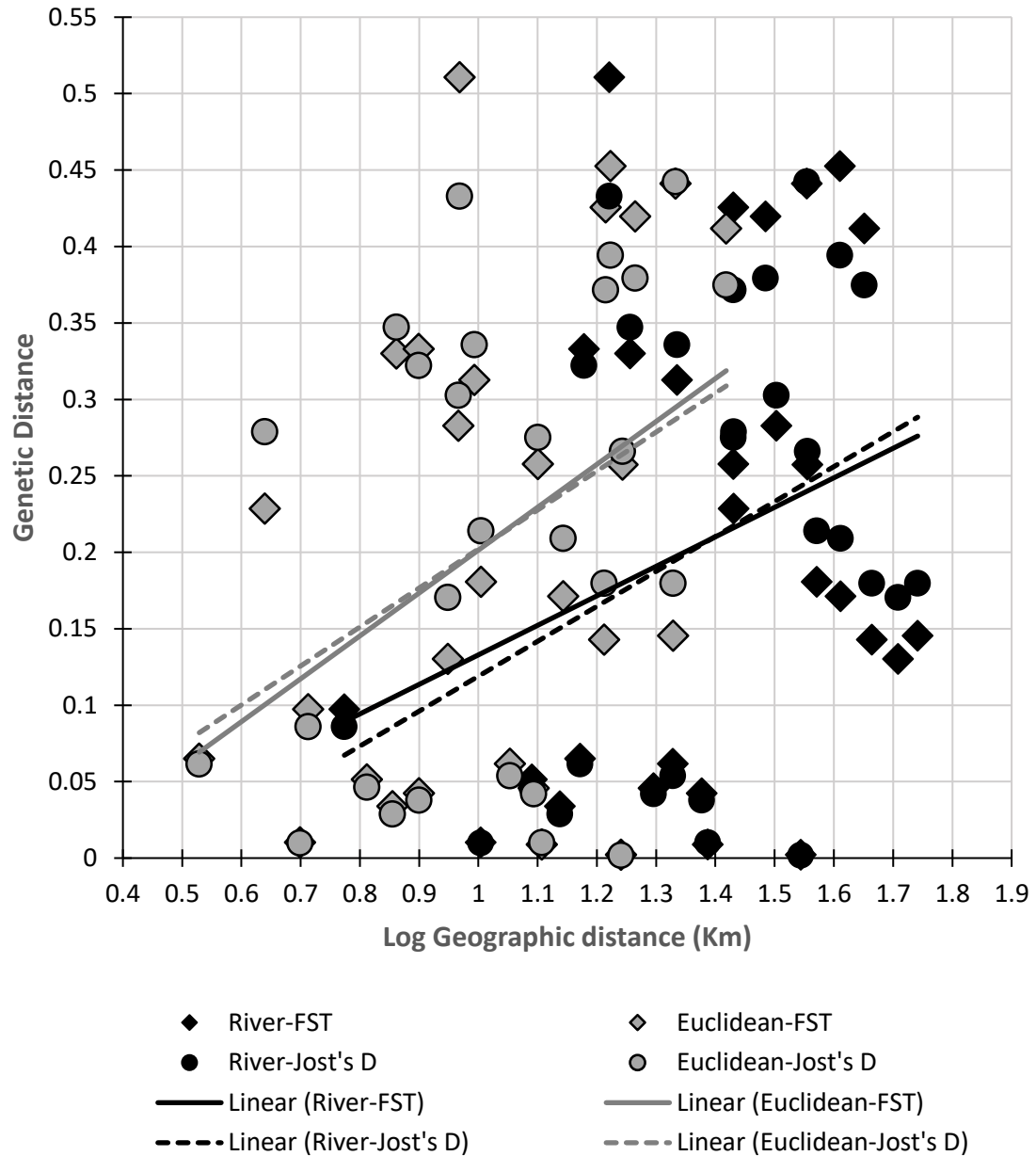


**Figure 4.** STRUCTURE analysis based on 19 microsatellite loci from 165 individuals across eight localities representing the Caney Fork form of *Cambarus pristinus*. a) Geographic distribution of the six population clusters inferred from the hierarchical STRUCTURE analyses (by colored polygons). Colors correspond to the inferred ancestral population depicted in the STRUCTURE plots (b-d). Letters by circles represent Site ID defined in Table 1. b) Initial STRUCTURE analysis with and without use of LOCPRIOR. c) Hierarchical STRUCUTRE plot using LOCPRIOR of Cluster A recovered in the initial STRUCTURE analysis. d) Hierarchical STRUCTURE plot using the LOCPRIOR for Cluster B recovered in the initial STRUCTURE analysis. b-d) Boxes designate localities and the color of each individual bar within a box denotes population ancestry of each individual. Letters below boxes and by circles correspond to site ID in Table 1. Delta K method is denoted by " $\Delta K$ "; the mean log-likelihood method is denoted by " $\mu L(K)$ ", and K represents the number of clusters inferred. An "\*" identifies the 6 distinct clusters recovered from all STRUCTURE runs.



**Figure 5.** Discriminant analysis of principle components (DAPC) based on 19 microsatellite loci and with individuals labeled by locality for the Caney Fork form of *Cambarus pristinus*. Ellipses denote the 6 genetic clusters recovered that correspond to the 6 clusters recovered in the STRUCTURE analyses (Figure 4). The top-right inset shows retained discriminant eigenvalues of the 5 principal components in relative magnitude. The first two principal axes are shown, with the first on the vertical axis.





**Figure 6.** Isolation by distance analysis for the eight sites of *Cambarus pristinus* showing the relationship between geographic (log-transformed river and Euclidean kilometers) and genetic distances ( $F_{ST}$  and Jost's D) based on 19 microsatellite loci. For corresponding pairwise values refer to Table 3. Comparisons using log-transformed river distances are in black and those using log-transformed Euclidean distances are in gray (River-  $F_{ST}$ :  $R^2=0.0776$ ;  $p=0.1479$ , Jost's D:  $R^2=0.1266$ ;  $p=0.091$ ; Euclidean-  $F_{ST}$ :  $R^2=0.1570$ ;  $p=0.0542$ , Jost's D:  $R^2=0.1511$ ;  $p=0.0541$ ).



## Literature Cited

- Adamack AT, Gruber B (2014) PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol Evol* 5:384–387. doi: 10.1111/2041-210X.12158
- Ahn D-H, Park M-H, Jung J-H, et al (2011) Isolation and characterization of microsatellite loci in the Korean crayfish, *Cambaroides similis* and application to natural population analysis. *Animal Cells Syst* (Seoul) 15:37–43. doi: 10.1080/19768354.2011.555137
- Allert AL, DiStefano RJ, Fairchild JF, et al (2013) Effects of historical lead-zinc mining on riffle-dwelling benthic fish and crayfish in the Big River of southeastern Missouri, USA. *Ecotoxicology* 22:506–521. doi: 10.1007/s10646-013-1043-3
- Allert AL, DiStefano RJ, Schmitt C, et al (2012) Effects of mining-derived metals on riffle-dwelling crayfish in southwestern Missouri and southeastern Kansas, USA. *Arch Environ Contam Toxicol* 63:563–573. doi: 10.1007/s00244-012-9797-9
- Alp M, Keller I, Westram AM, Robinson C (2012) How river structure and biological traits influence gene flow: a population genetic study of two stream invertebrates with differing dispersal abilities. *Freshw Biol* 57:969–981. doi: 10.1111/j.1365-2427.2012.02758.x
- Angermeier PL (1995) Ecological attributes of extinction-prone species: Loss of freshwater fishes of Virginia. *Conserv Biol* 9:143–158. doi: [10.1046/j.1523-1739.1995.09010143.x](https://doi.org/10.1046/j.1523-1739.1995.09010143.x)
- Anthony DM, Granger DE (2007a) An empirical stream power formulation for knickpoint retreat in Appalachian Plateau fluviokarst. *J Hydrol* 343:117–126. doi: 10.1016/j.jhydrol.2007.06.013
- Anthony DM, Granger DE (2007b) A new chronology for the age of Appalachian erosional surfaces determined by cosmogenic nuclides in cave sediments. *Earth Surf Proc Land* 32:874–887. doi: 10.1002/esp.1446
- Anthony DM, Granger DE (2004) A late tertiary origin for multilevel caves along the western escarpment of the Cumberland Plateau, Tennessee and Kentucky, established by cosmogenic <sup>26</sup>Al and <sup>10</sup>Be. *J Cave Karst Stud* 66:46–55
- Azuma N, Usio N, Korenaga T, et al (2011) Genetic population structure of the invasive signal crayfish *Pacifastacus leniusculus* in Japan inferred from newly developed microsatellite markers. *Plankt Benthos Res* 6:187–194. doi: [10.3800/pbr.6.187](https://doi.org/10.3800/pbr.6.187)
- Balkenhol N, Gugerli F, Cushman SA, et al (2009) Identifying future research needs in landscape genetics: where to from here? *Landsc Ecol* 24:455–463. doi: 10.1007/s10980-009-9334-z

- Barbaresi S, Fani R, Gherardi F, et al (2003) Genetic variability in European populations of an invasive American crayfish: Preliminary results. *Biol Invasions*. doi: 10.1023/A:1026133519707
- Barbaresi S, Santini G, Tricarico E, Gherardi F (2004) Ranging behaviour of the invasive crayfish, *Procambarus clarkii* (Girard). *J Nat Hist* 38:2821–2832. doi: 10.1080/00222930410001663308
- Belkhir K, Borsa P, Chikhi L, et al GENETIX 4.05, logiciel sous Windows TM pour la genetique des populations
- Berendzen PB, Simons AM, Wood RM (2003) Phylogeography of the northern hogsucker, *Hypentelium nigricans* (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. *J Biogeogr* 30:1139–1152. doi: 10.1046/j.1365-2699.2003.00888.x
- Berendzen PB, Simons AM, Wood RM, et al (2008) Recovering cryptic diversity and ancient drainage patterns in eastern North America: Historical biogeography of the *Notropis rubellus* species group (Teleostei: Cypriniformes). *Mol Phylogenet Evol* 46:721–737. doi: 10.1016/j.ympev.2007.07.008
- Black TR, Mattingly HT, Smith DD (2015) Utilization of pit telemetry to assess microhabitat affinities of stream-dwelling female crayfish during reproductive seclusion. *Freshw Crayfish* 21:71–82
- Bláha M, Žurovcová M, Kouba A, et al (2016) Founder event and its effect on genetic variation in translocated populations of noble crayfish (*Astacus astacus*). *J Appl Genet* 57:99–106. doi: 10.1007/s13353-015-0296-3
- Bohonak AJ (2002) IBD (Isolation by Distance): a program for analyses of isolation by distance. *J Hered* 93:153–154. doi: [10.1093/jhered/93.2.153](https://doi.org/10.1093/jhered/93.2.153)
- Bouchard R (1975) Geography and ecology of crayfishes of the Cumberland Plateau and Cumberland Mountains, Kentucky, Virginia, Tennessee, Georgia, and Alabama Part 2, the genera *Fallicambarus* and *Cambarus*. *Freshw Crayfish* 2:585–605. doi: 10.5869/fc.1975.v2.585
- Bubb DH, Thom TJ, Lucas MC (2006) Movement, dispersal and refuge use of co-occurring introduced and native crayfish. *Freshw Biol* 51:1359–1368. doi: 10.1111/j.1365-2427.2006.01578.x
- Buhay J (2009) “COI-like” sequences are becoming problematic in molecular systematic and DNA barcoding studies. *J Crustac Biol* 29:96–110. doi: 10.1651/08-3020.1
- Buhay J, Moni G, Mann N, Crandall K (2007) Molecular taxonomy in the dark: Evolutionary history, phylogeography, and diversity of cave crayfish in the subgenus

*Avitacambarus*, genus *Cambarus*. Mol Phylogenet Evol 42:435–448. doi: 10.1016/j.ympev.2006.07.014

Buhay JE, Crandall KA (2005) Subterranean phylogeography of freshwater crayfishes shows extensive gene flow and surprisingly large population sizes. Mol Ecol. doi: 10.1111/j.1365-294X.2005.02755.x

Colegrave N, Ruxton GD (2003) Confidence intervals are a more useful complement to nonsignificant tests than are power calculations. Behav Ecol 14:446–447. doi: 10.1093/beheco/14.3.446

Cornell SJ, Tregenza T (2007) A new theory for the evolution of polyandry as a means of inbreeding avoidance. Proc R Soc B 274:2873–2879. doi: 10.1098/rspb.2007.0926

Crandall KA, Buhay J (2008) Global diversity of crayfish (Astacidae, Cambaridae, and Parastacidae-Decapoda) in freshwater. Hydrobiologia 595:295–301. doi: 10.1007/978-1-4020-8259-7\_32

Crandall KA, De Grave S (2017) An updated classification of the freshwater crayfishes (Decapoda: Astacidea) of the world, with a complete species list. J Crustac Biol 37:615–653. doi: 10.1093/jcobiol/rux070

Crandall KA, Owen CL, Bracken-Grissom H, Stern D (2015) A synthetic phylogeny of freshwater crayfish: insights for conservation. Phil Trans R Soc B 370:20140009. doi: 10.1098/rstb.2014.0009

Creed R, Reed J (2004) Ecosystem engineering by crayfish in a headwater stream community. J North Am Benthol Soc J N Am Benthol Soc 23:224–236. doi: 10.1899/0887-3593(2004)023<0224:eebcia>2.0.co;2

Creed R (1994) Direct and indirect effects of crayfish grazing in a stream community. Ecology 75:2091–2103. doi: 10.2307/1941613

Davis CD, Epps CW, Flitcroft RL, Banks MA (2018) Refining and defining riverscape genetics: how rivers influence population genetic structure. WIREs Water 5:e1269. doi: 10.1002/wat2.1269

Davis DJ, Wieman AC, Berendzen PB (2014) The influence of historical and contemporary landscape variables on the spatial genetic structure of the rainbow darter (*Etheostoma caeruleum*) in tributaries of the upper Mississippi River. Conserv Genet 16:167. doi: 10.1007/s10592-014-0649-1

Di Rienzo A, Peterson AC, Garzat JC, et al (1994) Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci 91. doi: 10.1073/pnas.91.8.3166

- Di Stefano RJ, Black TR, Herleth-King SS, et al (2013) Life histories of two populations of the imperiled crayfish *Orconectes (Procericambarus) williamsi* (Decapoda: Cambaridae) in outhwestern Missouri, U.S.A. *J Crustac Biol* 33:15–24. doi: 10.1163/1937240X-00002109
- DiStefano RJ, Magoulick DD, Imhoff EM, Larson ER (2009) Imperiled crayfishes use hyporheic zone during seasonal drying of an intermittent stream. *J North Am Benthol Soc* 28:142–152. doi: 10.1899/08-072.1
- Do C, Waples RS, Peel D, et al (2014) NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size ( $N_e$ ) from genetic data. *Mol Ecol Resour* 14:. doi: 10.1111/1755-0998.12157
- Dorn NJ, Wojdak JM (2004) The role of omnivorous crayfish in littoral communities. *Oecologia* 140:150–159. doi; [10.1007/s00442-004-1548-9](https://doi.org/10.1007/s00442-004-1548-9)
- Dyer JJ, Brewer SK (2018) Habitat associations of three crayfish endemic to the Ouachita Mountain Ecoregion. *Southeaster Nat* 17:257–269. doi; [10.1656/058.017.0207](https://doi.org/10.1656/058.017.0207)
- Dyer JJ, Brewer SK, Worthington TA, Bergey EA (2013) The influence of coarse-scale environmental features on current and predicted future distributions of narrow-range endemic crayfish populations. *Freshw Biol* 58:1071–1088. doi: 10.1111/fwb.12109
- Earl D, VonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. doi; [10.1007/s12686-011-9548-7](https://doi.org/10.1007/s12686-011-9548-7)
- Eckert CG, Samis KE, Loughheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol Ecol* 17:1170–1188. doi: 10.1111/j.1365-294X.2007.03659.x
- El-Matbouli M, Soliman H (2006) Molecular diagnostic methods for detection of *Thelohania contejeani* (Microsporidia), the causative agent of Porcelain Disease in crayfish. *Dis Aquat Organ* 69:205–211. doi: 10.3354/dao069205
- Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Mol Ecol* 24:6021–6040. doi: 10.1111/mec.13454
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. doi; [10.1111/j.1365-294x.2005.02553.x](https://doi.org/10.1111/j.1365-294x.2005.02553.x)
- Fagan WF (2002) Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology* 83:3243–3249. doi: 10.1890/0012-9658(2002)083[3243:CFAERI]2.0.CO;2

- Fetzner JW, Crandall K a (2003) Linear habitats and the nested clade analysis: an empirical evaluation of geographic versus river distances using an Ozark crayfish (Decapoda: Cambaridae). *Evolution* 57:2101–18. doi: 10.1111/j.0014-3820.2003.tb00388.x
- Figiel CRJ (2016) Crayfish population genetic studies: assessment and trends. *Freshw Crayfish* 22:19–42. doi: 10.5869/fc.2016.v22-1.19
- Finlay JB, Buhay JE, Crandall KA (2006) Surface to subsurface freshwater connections: phylogeographic and habitat analyses of *Cambarus tenebrosus*, a facultative cave-dwelling crayfish. *Anim Conserv* 1:375–387. doi: 10.1111/j.1469-1795.2006.00046.x
- Finn DS, Blouin MS, Lytle DA (2007) Population genetic structure reveals terrestrial affinities for a headwater stream insect. *Freshw Biol* 52:1881–1897. doi: 10.1111/j.1365-2427.2007.01813.x
- Fluker BL, Kuhajda BR, Harris PM (2014) The influence of life-history strategy on genetic differentiation and lineage divergence in darters (Percidae: Etheostominae). *Evolution* (N Y) 68:3199–3216. doi: 10.1111/evo.12505
- Folmer O, Black M, Hoeh W, et al (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3:294–299
- Foster HR, Keller TA (2011) Flow in culverts as a potential mechanism of stream fragmentation for native and nonindigenous crayfish species. *J North Am Benthol Soc* 30:1129–1137. doi: 10.1899/10-096.1
- Frankham R (1995a) Conservation Genetics. *Annu Rev Genet* 29:305–327. doi: 10.1146/annurev.genet.29.1.305
- Frankham R (2005) Genetics and extinction. *Biol Conserv* 126:131–140. doi: 10.1016/j.biocon.2005.05.002
- Frankham R (1995b) Effective population size/adult population size ratios in wildlife: A review. *Genet Res (Camb)* 89:95–107. doi: 10.1017/S0016672308009695
- Frankham R, Bradshaw CJA, Brook BW (2014) Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biol Conserv* 170:56–63. doi: [10.1016/j.biocon.2013.12.036](https://doi.org/10.1016/j.biocon.2013.12.036)
- Frankham R, Lees K, Montgomery ME, et al (1999) Do population size bottlenecks reduce evolutionary potential? *Anim Conserv* 2:255–260. doi: 10.1017/S136794309900058X

- Furse JM, Wild CH, Villamar NN (2004) In-stream and terrestrial movements of *Euastacus sulcatus* in the Gold Coast Hinterland: developing and testing a method of assessing freshwater crayfish movements. *Freshw Crayfish* 14:213–220
- Gangloff MM, Perkins M, Blum PW, Walker C (2015) Effects of coal mining, forestry, and road construction on southern Appalachian stream invertebrates and habitats. *Environ Manage* 55:702–714. doi: 10.1007/s00267-014-0429-1
- Gaston KJ (1994) Geographic range sizes and trajectories to extinction. *Biodivers Lett* 2:163–170. doi: [10.2307/2999657](https://doi.org/10.2307/2999657)
- Gilbert KJ, Whitlock MC (2015) Evaluating methods for estimating local effective population size with and without migration. *Evolution (N Y)* 69:2154–2166. doi: 10.1111/evo.12713
- Glon M, Thoma R (2017) An observation of the use of Devil Crayfish (*Cambarus cf. diogenes*) burrows as brooding habitat by Eastern Cicada Killer Wasps (*Sphecius speciosus*). *Freshw Crayfish* 23:55–57
- Glon MG, Williams BW, Loughman Z (2019) *Lacunicambaru dalyae*: a new species of burrowing crayfish (Decapoda: Cambaridae) from the southeastern United States. *Zootaxa* 4683:361–380. doi: 10.11646/zootaxa.4683.3.3
- Gouin N, Grandjean F, Pain S, et al (2003) Origin and colonization history of the white-clawed crayfish, *Austropotamobius pallipes*, in Ireland. *Heredity (Edinb)* 91:70–77. doi: 10.1038/sj.hdy.6800282
- Gouin N, Grandjean F, Souty-Grosset C (2006) Population genetic structure of the endangered crayfish *Austropotamobius pallipes* in France based on microsatellite variation: biogeographical inferences and conservation implications. *Freshw Biol* 51:1369–1387. doi: 10.1111/j.1365-2427.2006.01570.x
- Gouin N, Souty-Grosset C, Bórquez J, et al (2011) Disentangling the impact of demographic factors on population differentiation of an endangered freshwater crayfish (*Austropotamobius pallipes*) using population density and microsatellite data. *Freshw Biol* 56:2105–2118. doi: 10.1111/j.1365-2427.2011.02629.x
- Grandjean F, Gouin N, Souty-Grosset C, Dieguez-Urbeondo J (2001) Drastic bottlenecks in the endangered crayfish species *Austropotamobius pallipes* in Spain and implications for its colonization history. *Heredity (Edinb)* 86:431–438. doi: [10.1046/j.1365-2540.2001.00849.x](https://doi.org/10.1046/j.1365-2540.2001.00849.x)
- Gross R, Palm S, Kõiv K, et al (2013) Microsatellite markers reveal clear geographic structuring among threatened noble crayfish (*Astacus astacus*) populations in northern and central Europe. *Conserv Genet* 14:809–821. doi: 10.1007/s10592-013-0476-9

- Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22:160–174. doi; [10.1007/bf02101694](https://doi.org/10.1007/bf02101694)
- Hedrick PW, Kalinowski ST (2000) Inbreeding depression in conservation biology. *Annu Rev Ecol Syst* 31:139–162
- Helms BS, Figiel C, Rivera J, et al (2013) Life-history observations, environmental associations, and soil preferences of the Piedmont Blue Burrower (*Cambarus [Depressicambarus] harti*) Hobbs. *Southeast Nat* 12:143–160. doi: 10.1656/058.012.0112
- Herrmann A, Schnabler A, Martens A (2018) Phenology of overland dispersal in the invasive crayfish *Faxonius immunitis* (Hagen) at the Upper Rhine River area. *Knowl Manag Aquat Ecosyst* 419:. doi: 10.1051/kmae/2018018
- Hoban S, Gaggiott O, Bertorelle G (2013) The number of markers and samples needed for detecting bottlenecks under realistic scenarios, with and without recovery: a simulation-based study. *Mol Ecol* 22:3444–3450. doi; [10.1111/mec.12258](https://doi.org/10.1111/mec.12258)
- Hollingsworth Jr. P, Near T (2009) Temporal patterns of diversification and microendemism in eastern highland endemic Barcheek Darters (Percidae: Etheostomatinae). *Evolution (N Y)* 63:228–243. doi; [10.1111/j.1558-5646.2008.00531.x](https://doi.org/10.1111/j.1558-5646.2008.00531.x)
- Hubisz M, Falush D, Stephens M, Pritchard J (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9:1322–1332. doi; [10.1111/j.1755-0998.2009.02591.x](https://doi.org/10.1111/j.1755-0998.2009.02591.x)
- Hughes JM, Bunn SE, Kingston DM, Hurwood DA (1995) Genetic differentiation and dispersal among populations of *Paratya australiensis* (Atyidae) in rainforest streams in outheast Queensland, Australia. *J North Am Benthol Soc* 14:158–173. doi; [10.2307/1467731](https://doi.org/10.2307/1467731)
- Hurry CR, Schmidt DJ, Hughes JM (2015) Phylogeography and limited distribution of the endangered freshwater crayfish, *Euastacus urospinosus*. *Aust J Zool* 63:236–244. doi: 10.1071/ZO15006
- Hurt C, Kuhajda B, Harman A, et al (2017) Genetic diversity and population structure in the Barrens Topminnow (*Fundulus julisia*): implications for conservation and management of a critically endangered species. *Conserv Genet* 1–12. doi: 10.1007/s10592-017-0984-0
- Imhoff EM, Mortimer RJG, Christmas M, Dunn AM (2009) Porcelain disease in white-clawed and signal crayfish in the UK. In: *Crayfish Conservation in the British Isles*. pp 49–56
- Janes JK, Miller JM, Dupuis JR, et al (2017) The K = 2 conundrum. *Mol. Eco* 26:3594–3602. Doi; [10.1111/mec.14187](https://doi.org/10.1111/mec.14187)

- Johansen JW (2018) Using multiscale rarity patterns to inform conservation actions in a diverse and understudied taxonomic group. Tennessee Technological University
- Johansen JW, Mattingly HT, Padgett MD (2016) Population Densities of Two Rare Crayfishes, *Cambarus obeyensis* and *Cambarus pristinus*, on the Cumberland Plateau in Tennessee. *Southeast Nat* 15:275–290. doi: 10.1656/058.015.0208
- Jombart T (2008) Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. doi: 10.1093/bioinformatics/btn129
- Jones S, Bergey E (2007) Habitat segregation in stream crayfishes: implications for conservation. *J North Am Benthol Soc* 26:134–144. doi: 10.1899/0887-3593(2007)26[134:HSISCI]2.0.CO;2
- Jost L (2008) G(ST) and its relatives do not measure differentiation. *Mol Ecol* 17:4015–26. doi: 10.1111/j.1365-294X.2008.03887.x
- Kahrl AF, Laushman RH, Roles AJ (2014) Evidence for multiple paternity in two species of *Orconectes* crayfish. *Can J Zool* 92:985–988. doi: 10.1139/cjz-2014-0132
- Kanno Y, Vokoun JC, Letcher BH (2011) Fine-scale population structure and riverscape genetics of brook trout (*Salvelinus fontinalis*) distributed continuously along headwater channel networks. *Mol Ecol* 20:3711–3729. doi: 10.1111/j.1365-294X.2011.05210.x
- Kerby JL, Riley SPD, Kats LB, Wilson P (2005) Barriers and flow as limiting factors in the spread of an invasive crayfish (*Procambarus clarkii*) in southern California streams. *Bio Con* 126:402–409 doi: 10.1016/j.biocon.2005.06.020
- Kessler RA, Black TR, Fraley SJ, Gangloff MM (2015) Molecular and morphological analyses reveal cryptic crayfish invasions in the Appalachian Mountains. *Freshw Crayfish* 21:51–62. doi: 10.5869/fc.2015.v21-1.51
- Koizumi I, Usio N, Kawai T, et al (2012) Loss of genetic diversity means loss of geological information: the endangered Japanese Crayfish exhibits remarkable historical footprints. *PLoS One* 7:e33986. doi: 10.1371/journal.pone.0033986
- Kopelman NM, Mayzel J, Jakobsson M, et al (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour* 15:1179–91. doi: 10.1111/1755-0998.12387
- Kozak KH, Blaine RA, Larson A (2006) Gene lineages and eastern North American palaeodrainage basins: phylogeography and speciation in salamanders of the *Eurycea bislineata* species complex. *Mol Ecol* 15:191–207. doi: 10.1111/j.1365-294X.2005.02757.x



- Kumar S, Stecher G, Tamura K, Dudley J (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33:1870–1874. doi: 10.1093/molbev/msw054
- Lamphere BA, Blum MJ (2011) Genetic estimates of population structure and dispersal in a benthic stream fish. *Ecol Freshw Fish* 21:75–86. doi: 10.1111/j.1600-0633.2011.00525.x
- Li Y, Guo X, Cao X, et al (2012) Population genetic structure and post-establishment dispersal patterns of the Red Swamp Crayfish, *Procambarus clarkia*, in China. *PLoS One* 7:40652. doi: 10.1371/journal.pone.0040652
- Light T (2003) Success and failure in a lotic crayfish invasion: the roles of hydrologic variability and habitat alteration. *Freshw Biol*. doi: 10.1046/j.1365-2427.2003.01122.x
- Lipták B, Mrugała A, Pekárik L, et al (2016) Expansion of the marbled crayfish in Slovakia: beginning of an invasion in the Danube catchment? *J Limnol* 75:305–312. doi: 10.4081/jlimnol.2016.1313
- Lodge DM, Kershner MW, Aloï JE, Covich AP (1994) Effects of an omnivorous crayfish (*Orconetes rusticus*) on a freshwater littoral food web. *Ecology* 75:1265–1281. doi: 10.2307/1937452
- Lodge DM, Taylor CA, Holdich DM, Skurdal J (2000) Nonindigenous crayfishes threaten North American freshwater biodiversity: lessons from Europe. *Fisheries* 25:7–20. doi: 10.1577/1548-8446(2000)025<0007:nctnaf>2.0.co;2
- Loughman ZJ, Skalican KT, Taylor ND (2013) Habitat selection and movement of *Cambarus chasmodactylus* (Decapoda:Cambaridae) assessed via radio telemetry. *Freshw Sci* 32:1288–1297. doi: 10.1899/12-166.1
- Loughman ZJ, Welsh SA, Sadecky NM, et al (2016) Environmental covariates associated with *Cambarus veteranus* Faxon, 1914 (Decapoda: Cambaridae), an imperiled Appalachian crayfish endemic to West Virginia, USA. *J Crustac Biol* 36:642–648. doi: 10.1163/1937240X-00002456
- Loughman ZJ, Williams BW (2018) *Cambarus polypilosus*, a new species of stream-dwelling crayfish (Decapoda: Cambaridae) from the Western Highland Rim of Tennessee, USA. 4403:171–185. doi: 10.11646/zootaxa.4403.1.10
- Lowe WH, Allendorf FW (2010) What can genetics tell us about population connectivity? *Mol Ecol* 19:3038–3051. doi: 10.1111/j.1365-294X.2010.04688.x
- Lowe WH, Kovach RP, Allendorf FW (2017) Population genetics and demography unite ecology and evolution. *Trends Ecol Evol* 32:141–152. doi: 10.1016/j.tree.2016.12.002

- Magoulick DD, DiStefano RJ (2007) Invasive crayfish *Orconectes neglectus* threatens native crayfishes in the Spring River Drainage of Arkansas and Missouri. *Southeast Nat* 6:141–150. doi: 10.1656/1528-7092(2007)6[141:ICONTN]2.0.CO;2
- Manangwa O, De Meeûs T, Grébaut P, et al (2019) Detecting Wahlund effects together with amplification problems: cryptic species, null alleles and short allele dominance in *Glossina pallidipes* populations from Tanzania. *Mol Ecol Resour* 19:757–772. doi: 10.1111/1755-0998.12989
- McGrath DA, Evans JP, Smith CK, et al (2004) Mapping land-use change and monitoring the impacts of hardwood-to-pine conversion on the southern Cumberland Plateau in Tennessee. *Earth Interact* 8:1–24. doi: 10.1175/1087-3562(2004)008<0001:mlcamt>2.0.co;2
- Meirmans PG, Hedrick PW (2011) Assessing population structure: FST and related measures. *Mol Ecol Resour* 11:5–18. doi: 10.1111/j.1755-0998.2010.02927.x
- Menger J, Unrein J, Woitow M, et al (2017) Weak evidence for fine-scale genetic spatial structure in three sedentary Amazonian understory birds. *J Ornithol* 159:355–366. doi: 10.1007/s10336-017-1507-y
- Micheletti SJ, Storfer A (2017) An approach for identifying cryptic barriers to gene flow that limit species' geographic ranges. *Mol Ecol* 26:490–504. doi: 10.1111/mec.13939
- Miller AD, Sweeney OF, Whiterod NS, et al (2014a) Critically low levels of genetic diversity in fragmented populations of the endangered Glenelg Spiny Freshwater Crayfish *Euastacus bispinosus*. *Endanger Species Res* 25:43–55. doi: 10.3354/esr00609
- Miller JM, Niraula BB, Reategui-Zirena EG, Stewart PM (2014b) Life history and physical observations of primary burrowing crayfish (Decapoda: Cambaridae) *Cambarus (Lacunicambarus) diogenes* and *Cambarus (Tubericambarus) polychromatus*. *J Crustac Biol* 34:15–24. doi: 10.1163/1937240X-00002199
- Moore JL (1985) Geologic map and mineral resources summary of the Pleasant Hill Quadrangle. Tennessee Div. Geol.
- Moore MJ, DiStefano RJ, Larson ER (2013) An assessment of life-history studies for USA and Canadian crayfishes: identifying biases and knowledge gaps to improve conservation and management. *Freshw Sci* 32:1276–1287. doi: 10.1899/12-158.1
- Moore PA, Bergman DA (2005) The smell of success and failure: The role of intrinsic and extrinsic chemical signals on the social behavior of crayfish. *Integr Comp Biol* 45:650–657. doi: 10.1093/icb/45.4.650
- Muncy L, Price SJ, Bonner SJ, Barton CD (2014) Mountaintop removal mining reduces stream salamander occupancy and richness in southeastern Kentucky (USA). *Biol Conserv* 180:115–121. doi: 10.1016/j.biocon.2014.09.040

- Murias dos Santo A, Cabezas MP, Tavares AI, et al (2016) tcsBU: a tool to extend TCS network layout and visualization. *Bioinformatics* 32:627–628. doi: [10.1093/bioinformatics/btv636](https://doi.org/10.1093/bioinformatics/btv636)
- Near TJ, Page LM, Mayden RL (2001) Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. *Mol Ecol* 10:2235–2240. doi: [10.1046/j.1365-294X.2001.01362.x](https://doi.org/10.1046/j.1365-294X.2001.01362.x)
- Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York
- Nguyen TTT, Austin CM, Meewan MM, et al (2004) Phylogeography of the freshwater crayfish *Cherax destructor* Clark (Parastacidae) in inland Australia: historical fragmentation and recent range expansion. *Biol J Linn Soc* 83:539–550. doi: [10.1111/j.1095-8312.2004.00410.x](https://doi.org/10.1111/j.1095-8312.2004.00410.x)
- Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Mol Ecol* 17:2258–2275. doi: [10.1111/j.1365-294X.2008.03750.x](https://doi.org/10.1111/j.1365-294X.2008.03750.x)
- Niemiller ML, Zigler KS (2013) Patterns of cave biodiversity and endemism in the Appalachians and Interior Plateau of Tennessee, USA. *PLoS One* 8:e64177. doi: [10.1371/journal.pone.0064177](https://doi.org/10.1371/journal.pone.0064177)
- Oliveira R, Pedro R•, Anastácio M (2015) Factors inducing overland movement of invasive crayfish (*Procambarus clarkii*) in a ricefield habitat. *Hydrobiologia* 746:135–146. doi: [10.1007/s10750-014-2052-9](https://doi.org/10.1007/s10750-014-2052-9)
- Oosterhout C Van, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538. doi: [10.1111/j.1471-8286.2004.00684.x](https://doi.org/10.1111/j.1471-8286.2004.00684.x)
- Parkyn S (2016) A review of current techniques for sampling freshwater crayfish. In: Kawai, T., Faulkes, Z., Scholtz G (ed) *Freshwater Crayfish: A Global Overview*. CRC Press, Boca Raton, pp 205–220
- Paz-Vinas I, Loot G, Stevens VM, Blanchet S (2015) Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Mol Ecol* 24:4586–4604. doi: [10.1111/mec.13345](https://doi.org/10.1111/mec.13345)
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539. doi: [10.1093/bioinformatics/bts460](https://doi.org/10.1093/bioinformatics/bts460)
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295. doi: [10.1111/j.1471-8286.2005.01155.x](https://doi.org/10.1111/j.1471-8286.2005.01155.x)

- Phillipsen IC, Kirk EH, Bogan MT, et al (2015) Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Mol Ecol* 24:54–69. doi: 10.1111/mec.13003
- Pilger T, Whitney JE, Turner TF (2015) Comparative conservation genetics of protected endemic fishes in an arid-land riverscape. *Conserv Genet*. doi: 10.1007/s10592-015-0707-3
- Pilger TJ, Gido KB, Propst DL, et al (2017) River network architecture, genetic effective size and distributional patterns predict differences in genetic structure across species in a dryland stream fish community. *Mol Ecol* 26:2687–2697. doi: 10.1111/mec.14079
- Pintor LM, Soluk DA (2006) Evaluating the non-consumptive, positive effects of a predator in the persistence of an endangered species. *Biol Conserv* 130:584–591. doi: 10.1016/j.biocon.2006.01.021
- Piry S, Luikart G, Cornuer J-M (1999) BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* 90:502–503. doi: [10.1093/jhered/90.4.502](https://doi.org/10.1093/jhered/90.4.502)
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Rambaut A (2012) FigTree v1.4
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution* (N Y) 43:223–225. doi: [10.2307/2409177](https://doi.org/10.2307/2409177)
- Richman NL, Bo'hm M, Adams SB, et al (2015) Multiple drivers of decline in the global status of freshwater crayfish (Decapoda: Astacidea). *Phil Trans R Soc B* 370:20140060
- Robinson CA, Thomy TJ, Lucas MC (2000) Ranging behaviour of a large freshwater invertebrate, the white-clawed crayfish *Austropotamobius pallipes*. *Freshw Biol* 44:509–521. doi: [10.1046/j.1365-2427.2000.00603.x](https://doi.org/10.1046/j.1365-2427.2000.00603.x)
- Robinson JD, Simmons JW, Williams AS, Moyer GR (2013) Population structure and genetic diversity in the endangered bluemask darter (*Etheostoma akatulo*). *Conserv Genet* 14:79–92. doi: 10.1007/s10592-012-0427-x
- Rohrbach G, Withers D (2006a) A Status Survey of the Caney Fork Crayfish (*Cambarus pristinus*) and Hardin County Crayfish (*Orconectes wrighti*). 1–81
- Rousset F (2008) GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Mol Ecol Resour* 8:103–106. doi: 10.1111/j.1471-8286.2007.01931.x

Schmidt B V., Schaefer J (2018a) Ecological and landscape effects on genetic distance in an assemblage of headwater fishes. *Ecol Freshw Fish* 27:617–631. doi: 10.1111/eff.12375

Schmidt B V., Schaefer JF (2018b) Comparative genetic isolation patterns for multiple headwater fishes in three geographic regions. *J Fish Biol* 92:1090–1109. doi: 10.1111/jfb.13570

Schorr MS, Backer JC, Schorra MS (2006) Localized effects of coal mine drainage on fish assemblages in a Cumberland Plateau stream in Tennessee. *J Freshw Ecol* 21:17–24. doi: 10.1080/02705060.2006.9664091

Schorr MS, Dyson MC, Nelson CH, et al (2013) Effects of stream acidification on lotic salamander assemblages in a coal-mined watershed in the Cumberland Plateau. *J Freshw Ecol* 28:339–353. doi: 10.1080/02705060.2013.778219

Serrao NR, Reid SM, Wilson CC (2018) Conservation genetics of Redside Dace (*Clinostomus elongatus*): phylogeography and contemporary spatial structure. *Conserv Genet* 3:409–424. doi: 10.1007/s10592-017-1012-0

Sexton JP, Hangartner SB, Hoffmann AA (2013) Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution (N Y)* 68:1–15. doi: 10.1111/evo.12258

Simmons JW, Fraley SJ (2010) Distribution, status, and life-history observations of crayfishes in western North Carolina. *Southeast Nat* 9:79–126. doi: 10.1656/058.009.s316

Song H, Buhay JE, Whiting MF, Crandall KA (2008) Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc Natl Acad Sci* 105:13486–13491. doi: [10.1073/pnas.0803076105](https://doi.org/10.1073/pnas.0803076105)

Taylor CA, Schuster GA (2004) *The Crayfishes of Kentucky, Illinois N*

Taylor CA, Schuster GA, Cooper JE, et al (2007) A reassessment of the conservation status of crayfishes of the United States and Canada after 10+ years of increased awareness. *Fisheries* 32:372–389. doi: 10.1577/1548-8446(2007)32

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. 111. Cladogram Estimation. *Genetics* 132:619–633

Thoma RF, Fetzner, Jr JW (2008) Taxonomic status of *Cambarus (Jugicambarus) jezerinaci*, Spiny Scale Crayfish (Powell River Crayfish). *Midwest Biodivers Inst*

- Thomas JR, Fisher J, Cable J, Griffiths SW (2018) Terrestrial dispersal of invasive signal crayfish during vulnerable life stages. *Behav Processes*. doi: 10.1016/j.beproc.2018.09.014
- Thornbury WD (1965) *Regional Geomorphology of the United States*. John Wiley and Sons, New York
- Tierney AJ, Godleski MS, Massanari JR (2000) Comparative analysis of agonistic behavior in four crayfish species. *J Crustac Biol* 20:54–66. doi: 10.1163/20021975-99990016
- Tonkin JD, Altermatt F, Finn DS, et al (2018) The role of dispersal in river network metacommunities: patterns, processes, and pathways. *Freshw Biol* 63:141–163. doi: 10.1111/fwb.13037
- Tregenza T, Wedell N (2000) Genetic compatibility, mate choice and patterns of parentage: Invited Review. *Mol Ecol* 9:1013–1027. doi: [10.1046/j.1365-294x.2000.00964.x](https://doi.org/10.1046/j.1365-294x.2000.00964.x)
- Turner TF, Trexler JC, Kuhn DN, Robison HW (1996) Life-history variation and comparative phylogeography of darters (Pisces: Percidae) from the North American Central Highlands. *Evol* 50:2023–2036. doi: 10.1111/j.1558-5646.1996.tb03589.x
- Vähä JP, Erkinaro J, Niemelä E, Primmer CR (2007) Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Mol Ecol* 16:2638–2654. doi: 10.1111/j.1365-294X.2007.03329.x
- Villanelli F, Cherardi F (1998) Breeding the crayfish, *Austropotamobius pallipes*: mating patterns, mate choice, and intermale competition. *Freshw Biol* 40:305–315. doi: [10.1046/j.1365-2427.1998.00355.x](https://doi.org/10.1046/j.1365-2427.1998.00355.x)
- Wagner MD, Blanton RE (2017) Do river drainage boundaries coincide with phylogeographic breaks in the Redband Darter? *Copeia* 105:696–711. doi: 10.1643/ci-16-414
- Walker D, Porter BA, Avise JC (2002) Genetic parentage assessment in the crayfish *Orconectes placidus*, a high-fecundity invertebrate with extended maternal brood care. *Mol Ecol* 11:2115–2122. doi: 10.1046/j.1365-294X.2002.01609.x
- Wang IJ, Bradburd GS (2014) Isolation by environment. *Mol. Ecol.*
- Wang J (2014) Estimation of migration rates from marker-based parentage analysis. *Mol Ecol* 23:3191–3213. doi: 10.1111/mec.12806
- Waples RS (2018) Null Alleles and  $F_{IS} \times F_{ST}$  Correlations. *J Hered* 457–461. doi: 10.1093/jhered/esy013

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci\*. *Conserv Genet* 7:167–184. doi: 10.1007/s10592-005-9100-y

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evol Appl* 3:244–262. doi: 10.1111/j.1752-4571.2009.00104.x

Ward RD, Woodwark M, Skibinski OF (1994) A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J Fish Biol* 44:213–232. doi: [10.1006/jfbi.1994.1021](https://doi.org/10.1006/jfbi.1994.1021)

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution (N Y)* 38:1358–1370. doi: [10.2307/2408641](https://doi.org/10.2307/2408641)

Welsh SA, Loughman ZJ (2015) Physical habitat and water quality correlates of crayfish distributions in a mined watershed. *Hydrobiologia* 745:85–96. doi: 10.1007/s10750-014-2095-y

Whiterod N, Zukowski S, Asmus M, et al (2016) Genetic analyses reveal limited dispersal and recovery potential in the large freshwater crayfish *Euastacus armatus* from the southern Murray-Darling Basin. *Mar Freshw Res* 68:213–225. doi: 10.1071/MF16006

Whitledge GW, Rabeni CF (1997) Energy sources and ecological role of crayfishes in an Ozark stream : insights from stable isotopes and gut analysis. *Can J Fish Aquat Sci* 54:2555–2563. doi: 10.1139/cjfas-54-11-2555

Williams CE, Bivens RD, Carter BD (2004) A crayfish survey of Wilkerson Creek and two of its tributaries Cumberland County Tennessee December 2004 Tennessee Wildlife Resources Agency

Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. *Conserv Genet* 6:551–562. doi: 10.1007/s10592-005-9009-5

Withers DI, McCoy RA (2005a) Distributional surveys for *Cambarus pristinus* and *Cambarus williamsi* , two endangered crayfish in Tennessee Final Report Contract # ID-05-0828-00 Tennessee Division of Natural Heritage Nashville , Tennessee. 1–57

Wright LL, Tregenza T, Hosken DJ (2008) Inbreeding, inbreeding depression and extinction. *Conserv Genet* 9:83. doi: 10.1007/s10592-007-9405-0

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

Yue GH, Li J, Bai Z, et al (2010a) Genetic diversity and population structure of the invasive alien Red Swamp Crayfish. *Biol Invasions* 12:2697–2706. doi: 10.1007/s10530-009-9675-1

Yue GH, Li JL, Wang CM, et al (2010b) High prevalence of multiple paternity in the invasive crayfish species *Procambarus clarkii*. Int J Biol Sci 6:107–115. doi: [10.7150/ijbs.6.107](https://doi.org/10.7150/ijbs.6.107)

Zachariah Peery M, Kirby R, Reid BN, et al (2012) Reliability of genetic bottleneck tests for detecting recent population declines. Mol Ecol 21:3403–3418. doi: [10.1111/j.1365-294X.2012.05635.x](https://doi.org/10.1111/j.1365-294X.2012.05635.x)

Zellmer AJ, Knowles LL (2009) Disentangling the effects of historic vs. contemporary landscape structure on population genetic divergence. Mol Ecol 18:3593–3602. doi: [10.1111/j.1365-294X.2009.04305.x](https://doi.org/10.1111/j.1365-294X.2009.04305.x)



## **Appendix A**

### **Supplementary Materials**

**Table S1.** Genetic diversity metrics summarized from the literature for several imperiled and non-imperiled crayfish and compared to those for *Cambarus pristinus* generated herein. Species are grouped by family with continent in parentheses. Metrics include IUCN conservation status (CS) with vulnerable (V), data deficient (DD), least concern (LC), and endangered (EN); mean alleles per locus ( $N_a$ ); allelic richness (AR); observed ( $H_o$ ) and expected ( $H_E$ ) heterozygosity. A “-“ indicates that genetic diversity metric was not reported.

Species	CS	$N_a$	AR	$H_o$	$H_E$	Citation
<b>Cambaridae (North America)</b>						
Pristine Crayfish ( <i>Cambarus pristinus</i> )	DD	6.10	2.66	0.41	0.42	
Red Swamp Crayfish ( <i>Procambarus clarkii</i> )	LC					
Invasive populations		6.42–14.33	-	0.56–0.81	0.70–0.89	(Li et al. 2012)
Native populations		14.58	-	0.65	0.88	(Li et al. 2012)
<b>Parastacidae (South America / Australia)</b>						
Glenelg Spiny Freshwater Crayfish ( <i>Euastacus bispinosus</i> )	V	1.0–2.3	1.07–1.75	0.004–0.42	0.03–0.36	(Miller et al. 2014a)
<b>Astacidae (Europe / North America)</b>						
Noble Crayfish ( <i>Astacus astacus</i> )	V	2.2–4.8	2.0–4.2	0.21–0.53	0.23–0.60	(Gross et al. 2013)
White-clawed Crayfish ( <i>Austropotamobius pallipes</i> )	EN	2.02	2.02	-	0.295	(Gouin et al. 2011)
Signal Crayfish ( <i>Pacifastacus leniusculus</i> )	LC	4 - 33	-	0.25–0.84	0.30–0.94	(Azuma et al. 2011)
<b>Cambaroididae (Asia)</b>						
Korean Crayfish ( <i>Cambaroides similis</i> )	DD	4.3–8.1	-	0.33–0.47	0.63–0.73	(Ahn et al. 2011)

**Table S2.** Primer sequences for 42 microsatellite loci optimized for the Caney Fork form of *Cambarus pristinus* and the Sequatchie form of *C. pristinus* including locus features: repeat motif, the number of alleles observed ( $N_a$ ), allele size range in base pairs (bp), and annealing temperature ( $T_a$ ). Data are summarized from the number of individuals indicated except for loci in bold which are summarized from our full data set of the Caney Fork form of *C. pristinus* with 169 individuals and an exploratory data set of the Sequatchie form with 21 individuals.

Primer	Primer sequence (5' – 3')	Repeat Motif	<i>C. pristinus</i> Caney Fork form (n=6)			<i>C. pristinus</i> Sequatchie form (n=1)		
			$N_a$	Allelic Range (bp)	$T_a$ (°C)	$N_a$	Allelic Range (bp)	$T_a$ (°C)
Cpris001	F: CACGACGTTGTAAAACGACGGCCTCCAACTTGAAGAGAC R: TCAACTCCGTGCCTACTCAC	AAAC	1	176	61	1	165	61
Cpris002	F: CACGACGTTGTAAAACGACGGCTTCCTGACTTCGACAAC R: TCTCCACTTCCCTGTAACCC	AAAC	1	226	61	1	217	61
Cpris004	F: CACGACGTTGTAAAACGACGGATTCCACCAGAGCTGAGAG R: AAAGCAATGACTCCTGCTGG	AAAC	1	285	63	1	285	63
<b>Cpris007</b>	F: CACGACGTTGTAAAACGACAGAAAGCCTCGCAACAAATG R: GGCTTCCTGGTCCCTATCATC	AAAG	5	328–344	61	1	332	61
Cpris008	F: CACGACGTTGTAAAACGACGTGTCCCGTCTTCAATTCCC R: AAGTCAGTGTGCACCTATG	AAAG	1	374	62	1	378	62
Cpris013	F: CACGACGTTGTAAAACGACTTAGCTTGCATCCCGAGTTG R: GACACCTGTTGCACACTACC	AAGT	1	309	62	1	309	62
<b>Cpris016</b>	F: CACGACGTTGTAAAACGACGAGCAGGGAAGATAGATGGAC R: TGCAGGGTGAACAGGTAAGG	ACAG	5	206–220	63	3	210–220	63

Primer	Primer sequence (5' – 3')	Repeat Motif	<i>C. pristinus</i> Caney Fork form (n=6)			<i>C. pristinus</i> Sequatchie form (n=1)		
			N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)	N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)
Cpris017	F: CACGACGTTGTAAAACGACAGGATCTGAACGAGTCTGTGG R: GCAGCCTCAAACCTTGGATAG	ACAG	1	244	62	1	240	62
Cpris018	F: CACGACGTTGTAAAACGACCTACGTCCGTCCCATACCC R: CGTTCCCGTGTAGTACTGAAG	ACAG	1	263	61	2	263–267	61
Cpris019	F: CACGACGTTGTAAAACGACATGTTGATAGGGTGAGGCTTC R: TCTGTGTCCGACTGTCTCTG	ACAG	2	273–277	62	1	271	61
<b>Cpris021</b>	F: CACGACGTTGTAAAACGACTCAGTGCTTGAGGAGAGAAGG R: TGCTTCTAATTAACCTGCCGG	ACAG	3	317–329	60	4	317–325	60
Cpris022	F: CACGACGTTGTAAAACGACTCTCATCTTAGATTCACTGCGG R: ACCTGTGTGTACCGATTGTG	ACAG	1	361	61	1	361	61
<b>Cpris026</b>	F: CACGACGTTGRAAAACGACTTCTCGCCATTTCTTCAGC R: GTGTGCCATTAAAGGAGCCTC	ACAT	7	232–256	62	6	236–264	62
<b>Cpris031</b>	F: CACGACGTTGTAAAACGACGGTATGCAGGCGRCAAGATG R: GACAGACTTCCATGCAAGGG	ACAT	6	322–338	62	4	322–334	62
<b>Cpris034</b>	F: CACGACGTTGTAAAACGACGCCCAACAGTCAGTCCAAAC R: ACTGGCAAGACAAATGAGCTG	ACAT	7	383–397	62	3	377–381	62
<b>Cpris037</b>	F: CACGACGTTGTAAAACGACTCATACCCATCCTGTGACTGG R: TGCCTCTGTTACCTGACAG	ACAT	7	422–454	63	3	426–438	63

Primer	Primer sequence (5' – 3')	Repeat Motif	<i>C. pristinus</i> Caney Fork form (n=6)			<i>C. pristinus</i> Sequatchie form (n=1)		
			N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)	N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)
Cpris039	F: CACGACGTTGTAAAACGACAGCAGCCTATGTCCATCTCG R: CGCCTGTCATCGATAACCTTG	ACGG	1	333	62	1	333	62
Cpris040	F: CACGACGTTGTAAAACGACACACACACAGTCAGTCGTCG R: GTACACTGCCTAGGACCAAAC	AGAT	2	141–149	63	2	141–149	63
Cpris043	F: CACGACGTTGTAAAACGACCACCCAGTTGTGCTTGCAG R: ACAGTCGGTCCATTTTCATCATC	AGAT	1	170	61	1	170	61
Cpris044	F: CACGACGTTGGTAAAACGACGATCGTTTGCAGTCAGAGGG R: TAGGCAGTGATGTGGTGGAG	AGAT	1	255	61	2	267–271	61
Cpris045	F: CACGACGTTGTAAAACGACGGTGGAGGATGGATTGGCAG R: TCGTAGATGTTGAATTAGCCGC	AGAT	1	297	62	1	304	62
<b>Cpris046</b>	F: CACGACGTTGTAAAACGACGATGAAGACCACTGCGCTAC R: CCGCCATCTATGTTGCTGTG	AGAT	12	303–347	61	9	295–339	61
Cpris047	F: CACGACGTTGTAAAACGACGGCCTCCCTCACTCTTACAG R: AACGAAACGCACCTCTTAGC	AGAT	2	331–335	62	1	335	62
<b>Cpris048</b>	F: CACGACGTTGTAAAACGACTCAGGGCACTCACTCTTAGC R: GGGTGGATGGATAGATGGAGAC	AGAT	4	374–382	61	2	374–378	61
<b>Cpris049</b>	F: CACGACGTTGTAAAACGACGGACTCGGTGCTGAACATTTTC R: TTGACACAGTTGGGAGGAGG	AGAT	7	440–468	62	4	448–464	62
Cpris055	F: CACGACGTTGTAAAACGACTCGGACCTTGCATAGTAGGC R: TCTAAACAGGAGGATGGGTTGG	ATCC	1	188	60	1	188	60

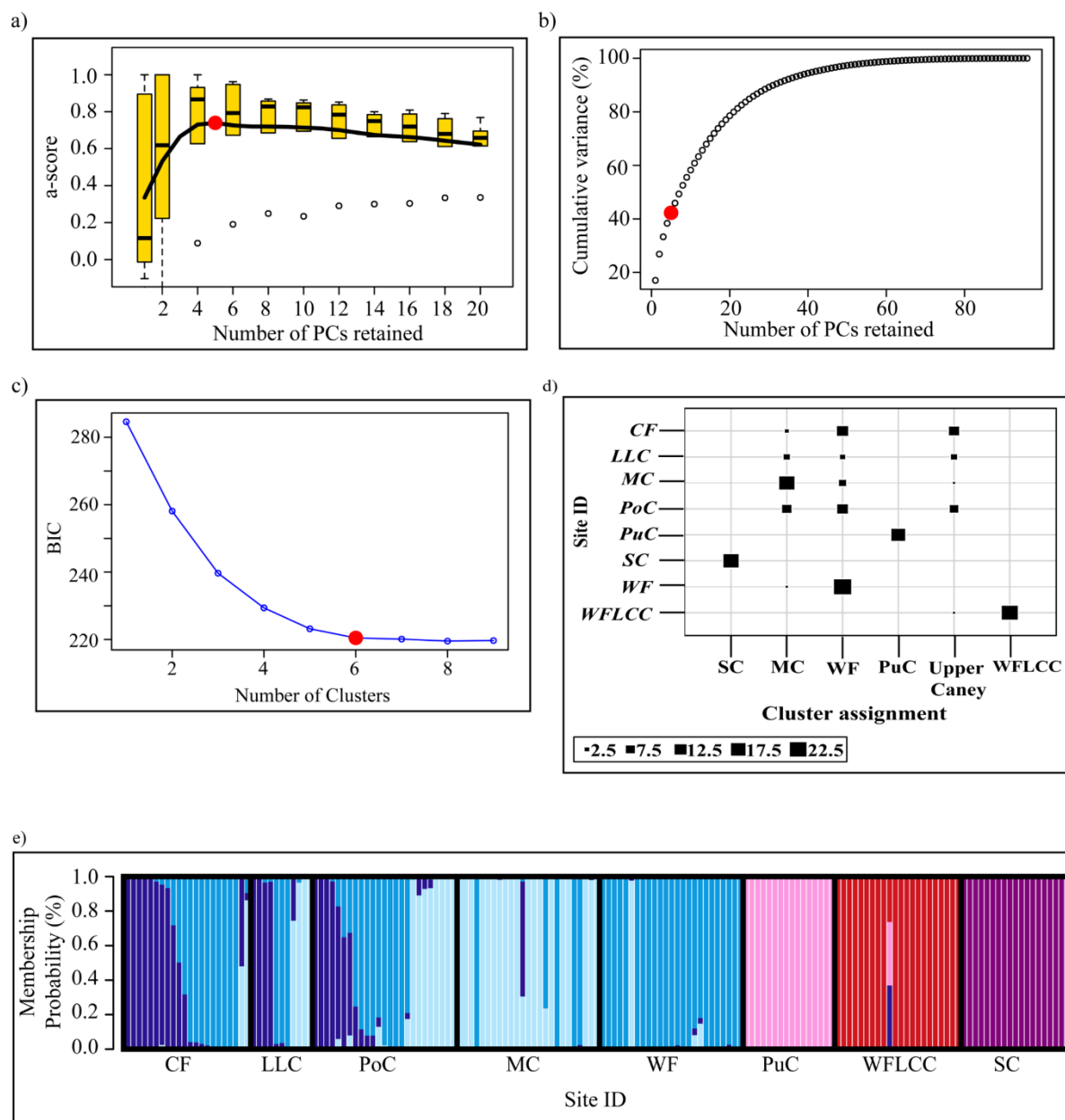
Primer	Primer sequence (5' – 3')	Repeat Motif	<i>C. pristinus</i> Caney Fork form (n=6)			<i>C. pristinus</i> Sequatchie form (n=1)		
			N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)	N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)
Cpris057	F: CACGACGTTGTAAAACGACTCTCTACCCATCCATCTGCG R: ACGCTAAGCTCCTCTGACTG	ATCC	1	348	62	1	348	62
<b>Cpris058</b>	F: CACGACGTTGTAAAACGACACACCTCCATACAACCAAC R: ACTCTCAGACTCTTAACGACCC	AAAC	2	162–164	62	2	164–168	62
<b>Cpris059</b>	F: CACGACGTTGTAAAACGACACAGGGAGTGGTAAGGTGTC R: GTTGGGTCAGTCTGGTGTG	AAAC	3	183–191	62	2	183–187	62
<b>Cpris061</b>	F: CACGACGTTGTAAAACGACTCCTGGGAAAGAAGTGTGTG R: TCACCACCTTTCCTCATCCC	AAAG	3	203–229	63	2	203–205	63
Cpris062	F: CACGACGTTGTAAAACGACGCCCTGTAACGCCTTTCAC R: CCATGGCGTTGTGGTAAGAC	AAAG	1	312	64	1	312	64
<b>Cpris063</b>	F: CACGACGTTGTAAAACGACTGTCGTCTCCCATTACCC R: TCAGGTGTTGACGAGGAAATC	ACAG	10	213–245	64	6	208–232	64
<b>Cpris064</b>	F: CACGACGTTGTAAAACGACCCCGCAAGCACAACCTAGG R: ACTACCGAGACCCTGAGATAG	ACAG	8	243–291	64	4	243–277	64
<b>Cpris066</b>	F: CACGACGTTGTAAAACGACATGGGACCAACTGAGACCTG R: GTTAACAGACCAACCAGCACG	ACAG	6	350–374	63	2	358–366	63
<b>Cpris068</b>	F: CACGACGTTGTAAAACGACGCCACCCAGAGTTCATGTG R: GGATTTACCTCCCAGATGG	ACAT	8	183–211	62	6	164–195	62
Cpris071	F: CACGACGTTGTAAAACGACAACCTCAAGAGGATGGGCTGG R: GCCAGTTTGATGTCTACCTTCG	ACAT	1	333	62	1	317	62

Primer	Primer sequence (5' – 3')	Repeat Motif	<i>C. pristinus</i> Caney Fork form (n=6)			<i>C. pristinus</i> Sequatchie form (n=1)		
			N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)	N <sub>a</sub>	Allelic Range (bp)	T <sub>a</sub> (°C)
Cpris073	F: CACGACGTTGTAAAACGACGAGAGCTGAGCCAAAGTGTG R: GCTTGTTACCTACTGCTGTC	ACAT	2	406–410	62	2	401–413	62
Cpris074	F: CACGACGTTGTAAAACGACTGACCAGCTTCCACGTATCC R: CCCATATTGACTGGCTTTTCGC	ACAT	2	362–368	64	2	362–368	64
<b>Cpris077</b>	F: CACGACGTTGTAAAACGACACTGCAAGTGTCTGGTTTCC R: TTTGTACAGTCGGTTTGGCC	AGAT	8	282–314	62	5	294–310	62
<b>Cpris078</b>	F: CACGACGTTGTAAAACGACACTGACAAACACCCATTCTCTC R: GACCTTCCTTCCACCACCC	AGAT	4	322–338	61	7	318–338	61
Cpris080	F: CACGACGTTGTAAAACGACGCCCATCAGTTGAACCATCG R: TGTAAGTGTCCCGCACTCC	ATCC	1	183	60	1	183	60
Cpris082	F: CACGACGTTGTAAAACGACACCACGACCAGCCTATGTTC R: AGACAGATGGGTGGATGGATAC	ATCC	1	380	62	1	380	62

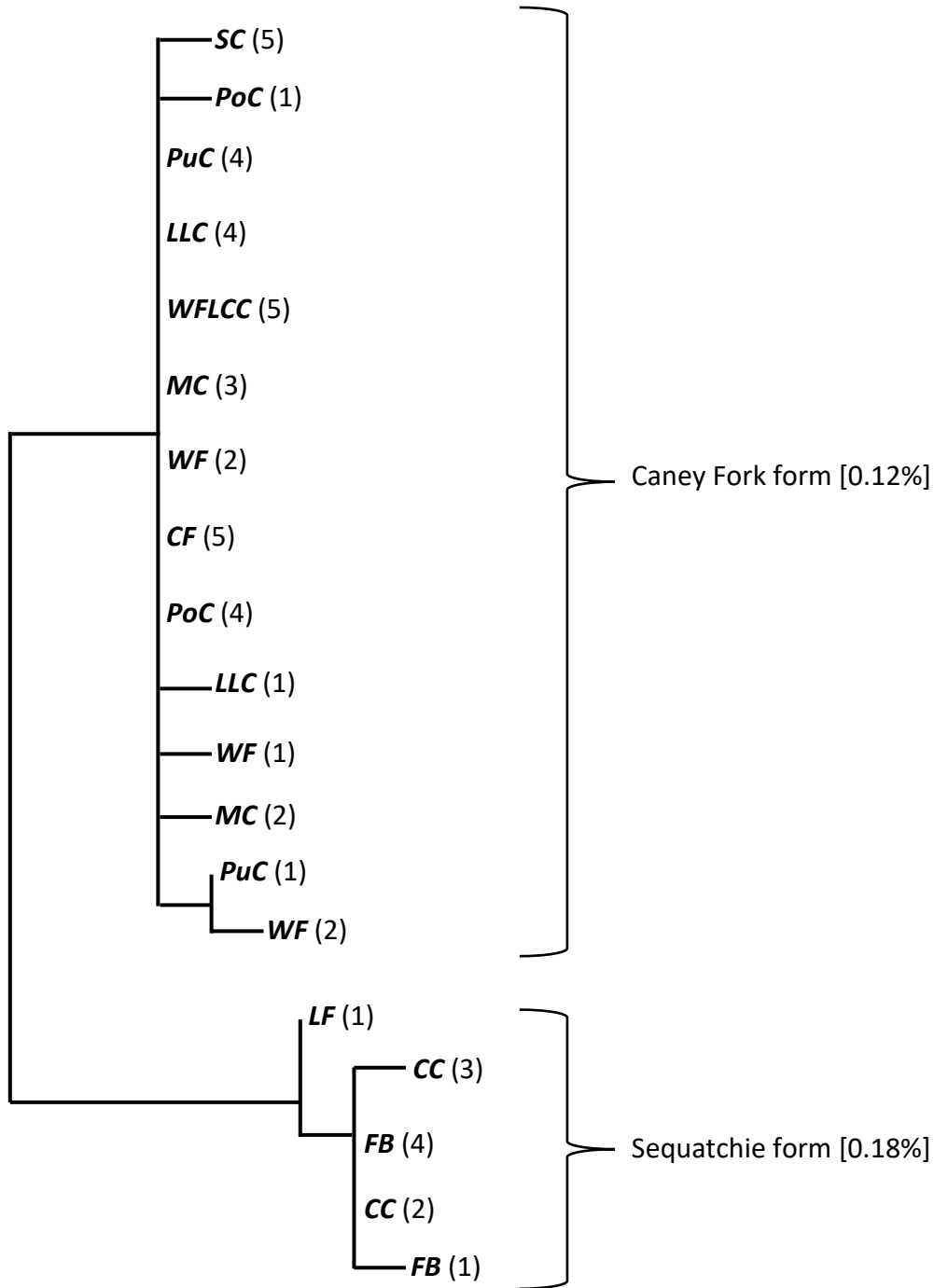
**Table S3.** Effective population size ( $N_e$ ) estimates generated herein and population census size ( $N_c$ ) estimates expressed as number of individuals/100 meters from Johansen et al. (2016). A “-” indicates that the metric was not reported. Site ID corresponds to Table 1.

Site ID	$N_e$ (CI)	$N_c$ (Regression method)	$N_c$ (mark-recapture method)
CF	<b>64.5</b> (27.7– $\infty$ )	<b>36</b>	<b>45</b>
LLC	$\infty$ (22.6– $\infty$ )	<b>5</b>	-
MC	<b>785.3</b> (53.1– $\infty$ )	<b>36</b>	-
PoC	<b>913</b> (56.0– $\infty$ )	<b>205</b>	<b>168</b>
WF	<b>1046</b> (58.3– $\infty$ )	<b>50</b>	<b>62</b>
WFLCC	<b>136.6</b> (52.7– $\infty$ )	<b>132</b>	<b>163</b>
Species overall	<b>15.5</b> (14.2–16.9)	<b>49 <math>\pm</math> 59</b>	-

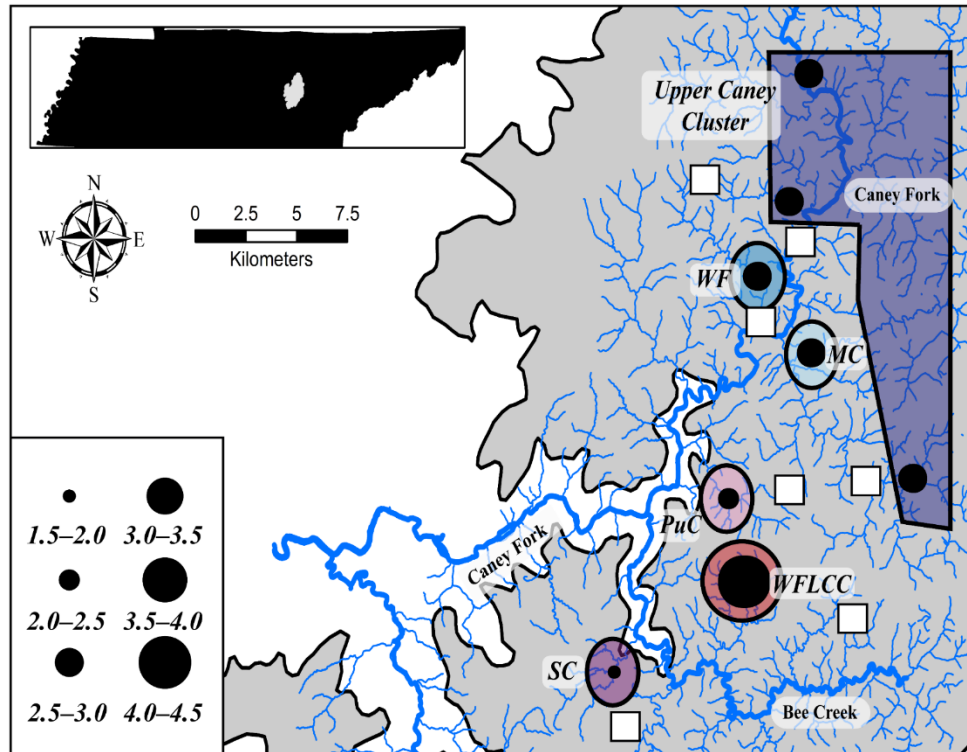




**Figure S1.** Criteria for selecting the number of retained principle components (PCs) and clusters for DAPC analysis. Red circles denote values retained in the DAPC analysis a) Graphic depicting recommended number of PCs to retain from the *optim.a* function in Adegenet. The a-score is the difference between the proportion of successful reassignment and values obtained using random groups retaining different numbers of PCs. b) Graph depicting number of retained PCs and explained cumulative variance. c) Graph depicting number of cluster and corresponding Bayesian information criterion (BIC) score. d) Proportion of individuals of a sampling site assigned to a genetic cluster denoted by DAPC (represented as the columns of boxes). The size of the squares represents the proportion of the individuals with the scale in the bottom left. Cluster letters correspond to Figure 5. e) Bar plot depicting membership probability of each individual to a cluster. Each box denotes a preassigned grouping by Site ID and color corresponds to cluster color used in Figure 4b-d. Each bar represents an individual and membership probability is determined by the 5 retained discriminant eigenvalues expressed as a percent.



**Figure S2.** Maximum likelihood tree recovered for *Cambarus pristinus* from the recovered mitochondrial COI gene showing within clade phylogenetic relationships. Support values >50% bootstrap support are shown. Site ID corresponds to Table 1 and numbers in parentheses represent number of individuals from that site recovered at that node. Numbers in brackets show within clade sequence divergence.



**Figure S3.** Geographic distribution of the six distinct populations of the Caney Fork form inferred from the STRUCTURE and DAPC analyses in relation to specific anthropogenic disturbances (strip mining and silviculture) in the region. Colored polygons represent a distinct population cluster and the black circle inside denotes a collected locality. The size of the black circle denotes the allelic richness recovered for each site and corresponds to the scale in the bottom left.