INFLUENCES OF HISTORICAL AND CONTEMPORARY ENVIRONMENTAL CONDITIONS ON THREATENED AND ENDEMIC AQUATIC ORGANISMS

by

Vaclav Alexei Sotola, M.S.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Aquatic Resources and Integrative Biology
May 2020

Committee Members:

Noland H. Martin, Chair

Timothy H. Bonner

Chris C. Nice

Daniel S. Stich

Jess W. Jones

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my major advisor, Dr. Noland Martin, and my (unofficial) co-advisor, Dr. Tim Bonner, for their support and encouragement throughout my graduate career at Texas State. Without their guidance, encouragement, and drive I would not have been able to achieve the success I have had here. You both have become close friends and colleagues, and I could not be more excited about my future as a researcher in this field. You have undoubtedly set me up for success! I would also like to thank the rest of my committee, Drs. Chris Nice, Daniel Stich, and Jess Jones for providing comments, suggestions, and encouragement throughout my time at Texas State University.

This project would not have been possible without funding provided by the Texas Comptroller of Public Accounts. Special thank you to all undergraduate and graduate students at Texas State University who helped with field and lab work, particularly Cody Craig, David Ruppel, Jeremy Maikoetter, Nicki Faucheux, Peter Pfaff, Marisa Quevedo, Christa Edwards, Sabrina Thiels, Danielle Terrell, Cameron Caldwell, Ryne Lehmann, Lauren Chappell, Rhys Woodruff, Anne Beckmann, and Erica Gomez. Our time working in the field and lab, having a blast collecting fishes and mussels, will never be forgotten. It's been tremendous, and I'm proud to say that I've definitely adopted Texas as a second home; I've been able to travel to pretty much every corner of this huge state with you all and I'll always cherish those memories. Additionally, thanks to BIO-West, Inc. researchers Brad Littrell, Kyle Sullivan, Jeff Jenkerson, and Jubentino Guajardo for their

help on the Colorado River mussel project. Crawling through the river on your hands and knees searching for mussels was fun with you all! Finally, thank you to all the folks in "Journal Club" for our weekly discussions and lab help, specifically Dr. Jim Ott, Alex Zalmat, Sarah Biliak, and Kate Bell. I think I learned more during our Thursday meetings than anywhere else while at Texas State! Cheers to everyone.

Outside of my professional help, I would not have been able to do this without my supporting cast of friends and family; they kept my head above the water when times were tough. I would need too many pages to list you all, so if you read this you know who you are. Thank you! Also, to the most important people in my life, my family: Janet, Vaclav, and Luke, thank you so much for helping me not only while at Texas State, but also throughout my whole life. No chance I am writing this dissertation without you three. Your unconditional support and love is appreciated and wholeheartedly returned.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

ABSTRACT

Historical and contemporary environmental conditions affect the distribution of aquatic species. The following three chapters seek to assess the degree to which conditions have impacted current distribution and population structure of fishes and mussels. The first chapter assesses genomic hybridization dynamics between the endemic and threatened *Macrhybopsis australis* and the widespread *M. hyostoma* in the Red River basin of Texas. This work found hybridization in a reach of the river upstream from Lake Texoma, and the distribution of hybrid and pure organisms is associated with several water quality parameters. The second chapter is a biogeographical assessment of the *Macrhybopsis* species complex within Texas. I used genomic techniques, and found a complex history of dispersal and vicariance, which likely occurred during the Pliocene and Pleistocene, that influenced current distributions. This work supports a stepping-stone model of dispersal, suggesting coastal drainages acted as islands, where species were able to transfer via connections during low sea levels. The third chapter assessed the effects of major floods on mussel populations using a closed robust mark-capture design to account for imperfect detection at two sites on the Colorado River. There were significant decreases in estimated abundances and apparent survival at one site, but not the other. The differential effects observed in estimated abundance and apparent survival among species were attributed to flood magnitude differences, substrate differences, and life-history traits of each species. These three chapters ultimately provide valuable information about the effects of historical and contemporary environmental conditions on

threatened and endemic aquatic organisms. I show that hybridization dynamics of fishes are significantly associated with water quality parameters, the evolutionary history of fishes in Texas shows patterns indicative of a stepping-stone model of dispersal, and population dynamics of freshwater mussels have a complex relationship with flood magnitude, substrate, and life-history traits. In summary, this work provides evidence that historical and contemporary environmental conditions shape the biology and ecology of aquatic organisms.

I. ASYMMETRIC INTROGRESSION BETWEEN FISHES IN THE RED RIVER BASIN OF TEXAS IS ASSOCIATED WITH VARIATION IN WATER QUALITY

Abstract

When ecologically divergent taxa encounter one another, hybrid zones can form when reproductive isolation is incomplete. The location of such hybrid zones can be influenced by environmental variables, and an ecological context can provide unique insights into the mechanisms by which species diverge and are maintained. Two ecologically differentiated species of small benthic fishes, the endemic and imperiled prairie chub, *Macrhybopsis australis*, and the shoal chub, *Macrhybopsis hyostoma*, are locally sympatric within the upper Red River Basin of Texas. I integrated population genomic data and environmental data to investigate species divergence and the maintenance of species boundaries in these two species. I found evidence of advanced-generation asymmetric hybridization and introgression, with shoal chub alleles introgressing more frequently into prairie chubs than the reciprocal. Using a Bayesian Genomic Cline framework, patterns of genomic introgression were revealed to be quite heterogeneous, yet shoal chub alleles were found to have likely selectively introgressed across species boundaries significantly more often than prairie chub alleles, potentially explaining some of the observed asymmetry in hybridization. These patterns were remarkably consistent across two sampled geographic regions of hybridization. Several environmental variables were found to significantly predict individual admixture, suggesting ecological isolation might maintain species boundaries.

Introduction

Speciation is usually not instantaneous (Grant 1981; Levin 1983, 2002; Wood et al. 2009), but rather involves the gradual buildup of reproductive isolation between diverging lineages (Coyne and Orr 2004). Understanding the genetic and ecological processes involved in the evolution of reproductive isolation is an important goal for evolutionary biologists, and hybrid zones have been used as "windows" into the speciation process (Harrison 1990). Hybridization occurs when divergent taxa meet and produce offspring of mixed ancestry and occurs in every major taxonomic group (Arnold, Sapir, & Martin, 2008). While hybridization can potentially be destructive to diversity at multiple scales, resulting in the erosion of genomic integrity, the fusion of taxa, or even the slow extirpation of taxa through genetic swamping (Allendorf et al., 2001; Edmands, 2007), hybridization can also act as an evolutionarily creative force leading to increased genetic diversity, adaptive introgression, and even hybrid speciation (Rieseberg et al. 2003; Martin et al. 2005, 2006; Arnold, Sapir, & Martin, 2008; Arnold and Martin 2009; Soltis and Soltis 2009; Gompert & Buerkle, 2016). Whether hybridization is the result of species naturally coming into secondary contact, anthropogenically-induced habitat modification, or introduction of closely-related species (Rhymer & Simberloff, 1996), researchers have begun to recognize the importance of identifying the degree to which hybridization and introgression are occurring at both genomic and ecological scales to better understand evolutionary processes and inform conservation decision-making.

The relatively recent ability to generate genome-wide data for non-model organisms has fortunately been accompanied with appropriate computational tools to process these exceptionally large datasets (Buerkle & Lexer, 2008; Gompert & Buerkle,

2009; 2011; 2012; Mandeville, Parchman, McDonald, & Buerkle, 2015). This has enabled evolutionary biologists to ask questions about the nature of reproductive isolation and introgressive hybridization at a genomic scale (Mandeville, Parchman, McDonald, & Buerkle, 2015; Sung et al., 2018). Genome-wide studies on non-model organisms have thus far provided strong support for the idea of a "genic view" of speciation (Wu, 2001), whereby the genomes of hybridizing species slowly accumulate loci that limit gene flow at very localized genomic scales and do not introgress, thus increasing reproductive isolation. However, introgression, either neutral or adaptive, can still occur throughout the remainder of the genome. Questions remain as to whether patterns of genomic isolation revealed in hybrid zones are consistent across divergent ecological contexts. The answers to these questions provide a context for predicting outcomes of hybridization and form the foundation for conservation management when hybridization involves species of concern.

In addition to an ecological effect, the direction of hybridization and introgression have been found to be influenced by the overall densities of the parental species simply because the rarer species have less opportunities to mate with conspecifics than with heterospecifics (i.e. the "Hubbs' effect"; Hubbs, 1955, Lepais et al., 2009). The causes of differences in densities has often been attributed to range shifts, mostly due to anthropogenic changes to the systems where these individuals occur (Perkin et al. 2015), although low amount of natural hybridization between sympatric species through natural range shifts have also been documented (Hasselman et al., 2014). Anthropogenic changes that can lead to hybridization, and a change in the local densities (or ranges) of species, include introductions of non-native species (or translocation of native species to new

watersheds), habitat fragmentation, habitat modification (Rhymer & Simberloff, 1996), or even purposeful fish stocking, resulting in a change in population densities (Marie, Bernatchez, & Garant, 2012; Heath, Bettles, & Roff, 2010; Lamaze et al., 2012). Additionally, changes in environmental variables, which can be altered via natural causes or anthropogenic disturbances, can increase the amount of hybridization between species (Marie, Bernatchez, & Garant 2012; Yau & Taylor, 2013). Previous research has found an association between increased amounts of hybridization with a decrease in the available habitat. Further, hybridization in fish has been found to be associated with a number of environmental factors (e.g. dissolved oxygen, temperature, and pH) that are believed to be population limiting factors (Marie, Bernatchez, & Garant 2012). Thus, understanding the ecological context in which hybridization is occurring is important in order to have a comprehensive understanding of the ecological drivers potentially influencing diversification and interspecific gene flow.

This study focused on two small benthic fishes, the shoal chub (*Macrhybopsis hyostoma*) and prairie chub (*Macrhybopsis australis*). These species are locally sympatric in the Red River along the border of Texas and Oklahoma upstream of Lake Texoma, an artificial reservoir created by the Denison Dam constructed in 1943 (Figure 1.1). The shoal chub has a broad distribution, occurring throughout the Mississippi River drainage and the West Gulf Slope drainages, including the lower reaches of Red River to upstream of the dam at Lake Texoma and the Brazos River (Eisenhour, 2004; Echelle et al., 2018). The prairie chub has a much more limited distribution and is endemic to the upper reaches of the Red River and several tributaries, including the Pease and Wichita rivers. The prairie chub currently is considered vulnerable and a species of greatest concern,

whereas the shoal chub is a species of least concern (Jelks et al., 2008; TPWD, 2012). It was previously assumed that meaningful introgression does not occur between sympatric *Macrhybopsis* within the Red River, based on morphological analysis (Eisenhour, 2004). However, allozyme data revealed that shoal chubs of the Red River are more genetically similar to the endemic prairie chub than shoal chubs elsewhere (Underwood et al., 2003; Echelle et al., 2018). This suggests that either the prairie chub is not a distinct taxon worth conservation consideration but rather a morphologically distinct subpopulation of the shoal chub, or that these two genetically distinct species are hybridizing, and the geographic and genomic extent of such hybridization is unknown.

Here I integrate population genomics and environmental data to investigate reproductive isolation and the maintenance of species boundaries between the shoal chub and prairie chub. Genotyping-by-sequencing (GBS) techniques were used to generate 39,122 SNPs which were then used to address three fundamental objectives: 1) quantify patterns of genetic structure and the geographic and genomic extent of hybridization, 2) examine patterns of excess ancestry at individual loci across two geographically disparate hybrid zones and identify the degree to which patterns of introgression were repeatable, and 3) determine the degree to which water quality parameters are associated with genetic structuring.

Methods

A total of 15 sites were sampled for genetic analysis: two from the Wichita River, four from the Pease River, six from the Red River upstream from Lake Texoma, one downstream from Lake Texoma, and two from the Brazos River (Figure 1.1; Table 1.1). Seins were utilized to collect shoal chubs and prairie chubs from these sites. Specimens

were euthanized with Tricaine Methanesulfonate (MS-222, Western Chemical, Inc.) and then subsequently stored in 95% ethanol.

Genomic DNA was extracted from fin clips taken from a total of 384 individuals in 96-well format using Qiagen DNEasy blood and tissue extraction kits and prepared for genotyping. For each sample, a reduced-complexity genomic library was prepared for GBS protocols modified from (Meyer & Kircher, 2010; Gompert et al., 2012; Parchman et al., 2012; Mandeville, Parchman, McDonald, & Buerkle, 2015). DNA from each individual was digested with the restriction enzymes EcoRI and MseI (New England Biolabs; NEB, Inc.). Fragments were labeled by ligating 8-10 base pair barcodes to the fragmented DNA. Two separate rounds of PCR were performed on these restriction-ligation products using Illumina primers, and the final PCR products were pooled into a single library. This library was then sent to the University of Texas Genomic Sequencing and Analysis Facility (Austin, TX, USA) and sequenced over two lanes on an Illumina HiSeq 2500 SR 150 platform after size selection between 300 and 400 base pairs via a Pippen Prep quantitative electrophoresis unit (Sage Science, Beverly, MA).

PhiX control sequences were identified by using Bowtie 2_db (Langmead & Salzberg, 2012). Raw reads that assembled to the PhiX genome were subsequently removed. A custom Perl script was used to remove Mse1 adapters and barcodes from sequence reads, correct single-base sequencing mutations in barcodes, and match sample IDs with unique barcode identifiers. Because a reference genome is not available for *Macrhybopsis*, a *de novo* assembly was performed using part of the dDocent variant calling pipeline (Puritz, Hollenbeck, & Gold, 2014). Specifically, unique reads were found for each individual, and reads with less than four copies and shared across less than

four individuals were filtered out of the dataset. The resulting filtered reads were assembled using CD-hit (Li & Godzik, 2006; Fu et al., 2012) with a threshold of 80% similarity. The scaffolds from this *de novo* assembly formed the basis of a reference-based assembly in which all sequence reads were assembled to the reference scaffolds using the *aln* and *samse* algorithms from BWA 0.7.5a-r405 (Li & Durbin 2009). SAMtools ver. 0.1.19 and BCFtools ver. 0.1.19 were used to identify variable sites (Single Nucleotide Polymorphisms – SNPs) and to calculate the Bayesian posterior probability that individual SNPs were variable (Li et al. 2009). In order for a locus to be included in the dataset, a minimum of 50% of all sampled fishes must have had at least one read at a particular locus (i.e. the "d" parameter in BCFtools was set at 0.5). For contigs containing more than one SNP, only a single randomly-chosen SNP was used for subsequent analyses. Importantly, individual SNP genotypes were not "called," but rather genotype likelihood estimates were assigned for each variable site for each individual. Furthermore, population allele frequencies were estimated directly from these genotype likelihood estimates, and SNPs with minor allele frequency of < 0.05 were excluded from the dataset. In all, genotype likelihood data were obtained for a total of 39,122 SNPs and used for population genomic analyses in this study.

To examine the genetic structuring of the shoal and prairie chubs, population genetic parameters were estimated using Entropy (Gompert et al., 2014a; Gompert et al., 2014b; Mandeville, Parchman, McDonald, & Buerkle, 2015). Entropy is a hierarchical model whereby an individual's assignment probability to each of any number of pre-assigned populations is estimated in a Bayesian framework. While interpretation of the output is similar to that of Structure (Pritchard, Stephens, & Donnelly, 2000; Falush,

Stephens, & Pritchard, 2003), Entropy accounts for variation in sequence coverage, sequence alignment, and genotyping errors, and produces posterior genotype probability distributions using prior probabilities from cluster allele frequencies (Gompert et al., 2014a). Models with different numbers of populations (k = 2 - 4) were compared; no attempts were made to identify the "best" k, but results of k = 2 - 4 runs are reported here, as an examination of all values of k all could provide a more comprehensive understanding of population structure. Posterior distributions of genotypes and admixture proportions were calculated for each k using Markov Chain Monte Carlo (MCMC) with 100,000 iterations sampling every 10th iteration. The first 5,000 iterations were discarded and each model for all k clusters was run twice. Calculation of the Gelman-Rubin diagnostic statistic and effective sample sizes were used to check chain convergence, and genotype and admixture proportions were subsequently averaged across both runs of each model. Posterior distributions for parameters were summarized as means, medians, and 95% credible intervals.

Population differentiation was explored by calculating pairwise Nei's G_{ST} (Nei, 1987). Allele frequencies were calculated in R (R Core Team, 2017) from the mean genotype posterior probabilities, which were in turn used to calculate pairwise G_{ST} values. In addition, population-level variation for each locality was reported using the genetic diversity index (π) calculated with SAMTools using the expectation-maximization (EM) algorithm, employing 20 iterations for each collection locale to achieve convergence of estimates (Li, 2011). In order to summarize the distribution of genetic variation, Principal Component Analysis (PCA) was performed in R on the genetic covariance matrix calculated from the genotype probability estimates generated in

Entropy (Gompert et al., 2014a).

A bidirectional stepwise regression was run to determine if the location of each species and hybrid individuals could be predicted by one or more environmental variables, which are known drivers of fish communities and can cause mortalities if they reach above or below tolerance levels (Barlow, 1958; Ostrand, 2000; Ostrand & Wilde 2001). Environmental data were collected during sampling events throughout the year as part of a larger project assessing the population dynamics and status of prairie chubs (Ruppel et al. 2017). Four environmental variables – specific conductance (µS/cm), pH, temperature (^{o}C), and dissolved oxygen (mg/L) were measured with a YSI 556 multi-probe sonde, and two additional variables, depth (m) and current velocity (m/s), were measured with a Marsh-McBirney Flo-mate model 2000 electromagnetic flowmeter. For the Red River basin collections, these environmental variables and river kilometer (rkm) were incorporated into a bidirectional stepwise multiple regression model to assess whether such variables could explain a significant portion of the variation in genetic assignment probabilities (q from k = 2) calculated from Entropy for the Red River basin fish only. Once terms deemed not useful by stepwise selection for use in the final model were removed, using the package *relaimpo*, relative importance of each environmental variable was assessed using the "lmg" type performing 10,000 bootstraps to determine confidence intervals for each variable's relative importance (Groemping, 2006). This analysis was done in R, and all data were log transformed prior to analysis.

While estimates of genome-wide admixture proportions can indicate whether or not hybridization is ongoing between divergent taxa, such estimates do not provide evidence as to how such admixture is occurring. Identifying admixture classes can

potentially provide a more detailed look at the long-term stability of hybrid zones and help to determine the extent to which both current hybridization and long-term introgression is occurring. For example, if most of the admixed individuals are found to be early-generation hybrids, this can indicate that either hybridization is a relatively recent phenomenon, or that late-generation hybrids may be largely unfit as they are not encountered. If, on the other hand, admixed individuals are shown to be of late-generation hybrids, this can indicate that the hybrid zone has been long established and that introgression across species boundaries is a possibility. Thus, I used an admixture class model in Entropy to estimate admixture classes (Q_{12}; Gompert et al., 2014a). This analysis assumes two source species; therefore, it was performed only on the Red River basin fishes due to the high degree of genetic differentiation between them and the Brazos River fish. I ran two independent MCMC analyses with 15000 iterations, sampled every 5^{th} iteration after a burn-in of 5000 iterations. Admixture classes were estimated from samples of both independent MCMC analyses.

The Bayesian genomic cline (BGC) model (Gompert & Buerkle, 2011; 2012) was used to quantify genome-wide variation in introgression among admixed individuals in two geographically separate areas. Because significant associations between water quality parameters and genetic assignment were found, these areas were run separately in an attempt to discover whether or not patterns of allelic introgression differed among divergent environmental conditions. The first location included hybrid individuals identified in the lower reach, just upstream from Lake Texoma at sampling locations RR_89, RR_81, RR_35, and RR_79 (downstream reach, N = 78), while the second included hybrid individuals identified from locations RR_283, PR_287, PR_6, and

WR_1919 (upstream reach, N = 16; see Results, Figure 1.1). Only individuals with Q_{12} (inter-population ancestry) values above 0.05 were included in the separate analyses. BGC is a hierarchical model that examines the probability of ancestry (ranging from 0 to 1) at individual loci as a function of an individual's hybrid index (*h*; also ranging from 0 to 1). Two locus-specific parameters were estimated, α and β. These reflect either an increase (+ α) or decrease (- α) in the probability of shoal chub ancestry for a locus relative to the probability of hybrid ancestry, while the parameter β specifies an increase (+ β) or decrease (-β) in the rate of change, with positive values indicating steeper clines and limited rates of introgression between species and negative values indicating wider clines with increased rates of introgression (Gompert & Buerkle, 2011; Gompert et al. 2012; Parchman et al., 2013). In order to estimate the marginal posterior probability distributions for α and β, two independent chains of MCMC were performed each with 50,000 iterations, sampled every 5th iteration, and following a 25,000 iteration burn-in. Outputs of the two chains were combined after determining both converged to the same stationary distributions. Medians and 95% CIs are reported for α and β, exceptional loci were identified as those where the 95% CIs of the parameter value did not intersect zero.

The degree to which exceptional α loci identified in the upstream reaches were also identified as exceptional in the downstream reaches was assessed by calculating the probability (*p*) that these two sets of loci were associated simply by chance using the following formula (from Sung et al. 2018):

$$\sum_{p=m}^{s} p = \frac{\begin{bmatrix} l \\ m \end{bmatrix} \times \begin{bmatrix} n-l \\ s-m \end{bmatrix}}{\begin{bmatrix} n \\ s \end{bmatrix}}$$

where *l* is the number of exceptional loci identified in the larger (downstream) hybrid

zone, *s* is the number of exceptional loci identified in the smaller (upstream) hybrid zone, *m* is the number of exceptional loci shared across both hybrid zones, and *n* is the total number of SNPs in the sample (e.g. 39,122 in the current study). This was only calculated for the α parameter as no exceptional β loci were found in the upstream hybrid zone.

Results

Sample sizes of fishes captured from the 15 sampling locales ranged from seven to 70 (Table 1.1). DNA sequencing resulted in a total of 409,970,747 reads with an average of 1,070,419 sequences per individual. Individuals with low coverage (mean of <2 reads per locus) were not included in analyses, resulting in a total of 368 individuals with an average of 7.01 (SD = 1.62) reads per locus per individual. Overall, a total of 39,122 SNPs were ultimately included in the analyses.

Genotype likelihood estimates were calculated for all SNPs for each individual. Highest π diversity levels were found in the Red River sites upstream from Lake Texoma, followed by the Wichita and Pease rivers, with the lowest observed in the Brazos River (Table 1.1). The highest amount of genetic differentiation (G_{ST}) was found between the Brazos River sites and all the Red River drainage sites, indicating higher relative genetic differentiation between shoal chubs of different drainages than shoal chubs and prairie chubs within the Red River drainage. Next highest levels of differentiation were found with the site downstream from Lake Texoma (RR_259) compared to the Pease and Wichita rivers sites. Lowest relative genetic differentiation was found among the Pease and Wichita River sites and among the mainstem Red River sites (Table 1.2).

Principal component I explained 39.76% of the variation, and principal component II explained 13.51% of the variation in the genotypic data (Figure 1.2). Three

primary clusters were indicated by PCA: (1) shoal chubs from the Brazos River, (2) prairie chubs from the Pease, Wichita, and upper Red River sites, and (3) shoal chubs from the downstream Red River sites. PC I segregated shoal chubs from the Brazos River sites from shoal and prairie chubs from the Red River drainage, while PC II segregated prairie chubs from the Pease and Wichita rivers sites from shoal chubs from the Red River and Brazos River sites. Hybridization between the shoal chub and prairie chub is evident from intermediate individuals observed from the mainstem Red River sites and Pease rivers sites with intermediate PC I and PC II scores (Figure 1.2).

Admixture proportions were calculated in Entropy for k = 2 – 4 with all of the sampled sites (Figure 1.3). Similar to the PCA results, for k = 2, the model separated individuals from the Brazos River and Red River drainages into two genetic clusters. Shoal chubs in the Red River show mixed ancestry between the two genetic clusters. At k = 3, individuals from the two Brazos River sites (Brazos River shoal chub) were separate from Red River sites. All the individuals sampled from the site downstream from Lake Texoma, RR_259, along with some individuals from sites upstream from Lake Texoma (RR_35, RR_89, RR_81, and RR_79) are grouped into the second (Red River shoal chub) cluster with very high probability. The third genetic cluster (prairie chubs) consisted of individuals from the Wichita, Pease, and upstream Red River sites (RR_283, RR_70). A majority of individuals from RR_79 downstream to the site just upstream of Lake Texoma (RR_35) had intermediate assignment probabilities between the Red River shoal chub and prairie chub genetic clusters, indicating a clear hybrid zone. For k = 4, genetic clusters resemble those observed in k = 3, although no individuals had 100% assignment probability to additional clusters that were added and do not appear to

provide any easily discernable biological interpretations.

Bidirectional stepwise selection for water quality variables predicting assignment probability from Entropy (q) found the model with the lowest AIC score included specific conductance, depth, dissolved oxygen, river kilometer, pH, and current velocity; temperature was removed by the stepwise procedure. The selected model was significant ($F_{6,302} = 108.80$, $P < 0.001$) and explained approximately 67.74% of the variation in q (Table 1.3). Of the explained variation in the final model, specific conductance (95% CI: $29.71 < 34.12 < 39.00$) explains the highest percentage. Specific conductance, current velocity, and pH were found to have a positive relationship with q, indicating that as these variables increase, prairie chub ancestry also increases; the opposite was found for dissolved oxygen, depth, and river kilometer.

Admixture class estimates (Q_{12}) revealed mixed ancestry for individuals from several sites upstream from Lake Texoma (Figure 1.4). Individuals collected from these sites had genomic regions that were inherited from two different source species (nonzero Q_{12}), indicative of hybridization. All of the individuals taken downstream from Lake Texoma (RR_259) were of pure shoal chub ancestry. A majority of individuals from the Pease River, upper Red River sites (RR_283 and RR_70), and the Wichita River were predominately of prairie chub ancestry, although a minority of individuals did have small amounts of mixed ancestry. No pure shoal chub was captured upstream of RR_81. A majority of the individuals from sites upstream of Lake Texoma (RR_35, RR_89, RR81, and RR_79) are not easily assignable to early-generation hybrids due to backcrossing between hybrid lineages and parental species (Nadeau 2014), yet one individual does appear to be genetically indistinguishable from an F_1 hybrid (individual nearing 0.5 for

genome-average ancestry0 and 1.0 for inter-population ancestry), indicating that while initial F_1 hybridization may be rare, those hybrids – as well as later-generation hybrids are likely fit.

In hybrid individuals captured in the upstream reach, the posterior estimates of genomic cline parameter α was variable across loci, ranging from -0.47 to 1.05 (Figure 1.5). The parameter β was less variable with posterior estimates ranging from -0.15 to 0.06. In total, 40 loci (0.10%) were found to be exceptional (95% credible intervals do not include zero), all of which had positive α values, meaning that these shoal chub alleles were likely selectively favored regardless of the genomic background in which they occurred. None of the loci in this upper reach were found to have exceptional β values.

In hybrid individuals from the downstream reaches, α was again highly variable across loci, ranging from -1.43 to 2.03, while measures of β ranged from -1.08 to 0.81. In total, 6,393 loci (16.34%) were found to have exceptional α values; 5,214 of those had positive α values (i.e. shoal chub alleles had higher frequency regardless of the genomic background), while 1,159 had negative α values (i.e. prairie chub alleles had higher frequency regardless of the genomic background; Figure 1.5C, D), two loci revealed exceptionally positive β values (i.e. were overrepresented in conspecific backgrounds, while underrepresented in heterospecific backgrounds – consistent with loci affecting reproductive isolation and are likely responsible for reproductive isolation between the two taxa), and 18 had exceptionally negative β values (indicating that bi-directional selective introgression is likely to occur for these loci). I found significant concordance ($P < 0.0001$) across the two hybrid zones. Of the 40 exceptional loci found in the

individuals sampled in the upstream reach, 35 loci were also identified as exceptional loci in individuals sampled in the downstream reach (Figure 1.5A, B). This is consistent with those loci imparting a selective advantage throughout the range where hybrids occur, with such selective advantages not being site-specific.

Discussion

Previous studies utilizing small genetic datasets and morphological data concluded that shoal chubs and prairie chubs were morphologically distinct species with no meaningful hybridization and introgression occurring (Eisenhour, 2004; Underwood et al., 2003; Echelle et al., 2018). The current genomic results corroborate that the two species are in fact distinct taxa, yet there is evidence of interspecific hybridization in the lower reaches of the river basin where they co-occur. Hybridization occurs predominately in the reach immediately upstream from Lake Texoma, with much less hybridization in the upper Red River and Pease River with trace amounts of admixture evident in the Wichita River. As in previous studies, I found higher genetic differentiation between two populations of shoal chubs in different drainages (Brazos River and Red River) than between shoal chubs and prairie chubs within the Red River (Underwood et al., 2003; Echelle et al., 2018). I found that collection locales consisting predominately of hybrid individuals had higher genetic diversities (π), which is expected if alleles from divergent taxa are contributing to allelic diversity at these collection locales (Zalapa, Brunet, & Guries, 2010). Additionally, there is asymmetry with regards to both the hybrid zone and introgression of alleles, with the hybridization being predominately dominated by shoal chubs' background and their alleles into largely prairie chub backgrounds. Attempts to classify the hybrid individuals into early-generation hybrid classes (e.g. F1, F2, or BC1

individuals) were largely unsuccessful, with only a single mixed-ancestry individual being categorized as a possible F1. All other hybrid individuals were late-generation hybrids that were not easily categorized into specific hybrid classes (Nadeau 2014), likely indicating that hybridization has persisted for many generations. This is likely because a majority of the sampled individuals were hybrids, with only a few pure individuals of each species, thus there is little potential for early generation hybridization to occur.

The hybrid zone has a broad pattern of asymmetry. Asymmetry in hybrids has been attributed to several factors including differences in generation time (Barton 1986), mating behavior (Lamp & Avise 1986; Konkle & Philipp 1992), fitness (Ostberg et al., 2004), or relative abundances of the parental species (Lepais et al. 2009). Among these factors, relative abundances of parental species differ within the hybrid zone, with shoal chubs being much more common than prairie chubs. In the reach where hybridization occurs, relative abundances of putatively identified shoal chubs (range 2.18 – 2.83%) are approximately 3 to 4 times higher than that of putatively identified prairie chubs (0.31 – 0.86%; Ruppel et al., 2017). This potentially results in pure prairie chub individuals and hybrids having more mating opportunities with heterospecifics than with conspecifics. Greater abundance of shoal chubs and lesser abundance of prairie chubs within the hybrid zone correspond with an overlap in distributions between prairie chubs and shoal chubs. The hybrid zone might be a natural secondary contact zone between the two species and therefore represent natural hybridization events between closely related taxa. Lake Texoma, which is located downstream of the hybrid zone, might have exacerbated hybridization between these two species. Dams disrupt the habitat and environmental heterogeneity of rivers, homogenizing habitats (Santucci et al. 2005),

which can lead to an increase in introgressive hybridization (Seehausen et al. 2008; Hasselman et al. 2014). Thus, it is possible that the construction of Lake Texoma could have restricted a species of mobile *Macrhybopsis* (Wilde, 2016; Worthington 2016) upstream and, along with altering upstream habitats (e.g., deeper water and more similar to habitats associated with Shoal Chub; Eisenhour 2004), could have anthropogenically inflated Shoal Chub numbers in the zone.

BGC analyses demonstrated that introgression rates were quite variable across the genome. This comports with other studies examining genome-wide rates of introgression in hybrid zones (Gompert et al., 2012; 2014; Kingston et al., 2017; Parchman et al., 2013; Payseur, 2010; Teeter et al., 2010; Yuri et al., 2009) including fish (Nolte, Gompert, & Buerkle, 2009; Schaefer, Duvernell, & Campbell, 2016). Of the 39,122 loci examined in the current study, 16% were revealed to have exceptional α values in the downstream reach, with shoal chub alleles having a > 5.5-fold chance of introgressing across species boundaries. The same pattern was observed in hybrids in the upper reach; while only 40 loci with exceptional α values were found in that hybrid zone, all of them were revealed to be crossing from shoal chubs to prairie chubs. The fact that such a small number of loci were found to have exceptional α values in individuals captured in the upper reach is not surprising given that the number of hybrid individuals used in this analysis was quite small (N = 16). Exceptional α values are consistent with selection or adaptive introgression, thus, it is possible that these alleles are selectively advantageous and have spread into the alternative genomic background (Whitney et al., 2006). This indicates that some of the asymmetric introgression observed here could be explained by the fact that shoal chub alleles are more often than not selectively advantageous. However, stochastic

evolutionary processes (i.e. drift) in small populations can also contribute to increased exceptional α values. It is difficult to know if the populations of chubs in the Red River are small and potentially influenced by drift, yet similar levels of genetic diversities and the fact that relative abundances have largely increased or remained stable since the 1940s in the areas sampled by this study (Ruppel et al., 2017) suggests drift should not be acting strongly on these fishes. Thus, selection or adaptive introgression seem the likely drivers of the high amount of exceptional α values in these fishes, perhaps some of which is due to several extrinsic factors throughout the basin.

In this study, of the 40 exceptional loci found in hybrid individuals captured in the upstream reach, 35 of them were also found to be exceptional in the downstream reach, and this overlap was greater than expected by chance. This likely indicates that the selective advantages afforded by these loci are not simply site-specific, but occur basin-wide, and these loci are strong candidates for having moved upstream into largely pure prairie chub populations via selection, especially as no pure shoal chub individuals are encountered in the area. This is in contrast to previous studies assessing multiple hybrid zones in fishes (Nolte, Gompert, & Buerkle, 2009; Aboim et al., 2010). These studies found differential patterns of introgression between two hybrid zones and attributed these different patterns to extrinsic factors that are differentially affecting the populations in different areas. In our study, shoal chub alleles were crossing more often into prairie chubs from both upstream and downstream reaches, and it is likely that at least some of these alleles are selectively advantageous regardless of where the individuals were spawned. It is unknown whether or not the hybrids analyzed in the upstream reaches actually spawned there or traveled from the downstream reaches, however most of the

hybrid individuals in the upstream reaches were late generation hybrids with largely prairie chub backgrounds.

The distributions of both species and their hybrids in the Red River are strongly associated with several environmental variables, which are known to be important factors in structuring many other fish communities (Barlow, 1958; Ostrand 2000; Ostrand & Wilde, 2001; 2004). Marie, Bernatchez, & Garant (2012) found both positive and negative associations between physiochemical environmental conditions (e.g. temperature, dissolved oxygen, and pH) and hybridization rates, and suggested these may be limiting factors on fishes that are affecting their ability to reproduce in certain areas. Here, I found several water quality variables that are significant predictors of their genetic assignment probability (q), including specific conductance, pH, current velocity, depth, and dissolved oxygen. I found that temperature was not a significant predictor of q, which is in contrast to other published studies of fish hybridization (Marie, Bernatchez, & Garant 2012; Yau & Taylor, 2013), indicating temperature does not influence hybridization. In particular, specific conductance explained a majority of the variation in the model, indicating it was the strongest environmental predictor of the genomic composition of individuals that was measured, and as such may be a limiting factor in these species' distributions. As specific conductance increases, one is more likely to encounter prairie chubs, whereas shoal chubs are more likely to be found in areas with lower specific conductance, and admixed individuals were captured more often in areas with intermediate specific conductance. This is an interesting association and future experimental studies specifically testing the overall fitness of prairie chubs, shoal chubs, and their hybrids at various water quality (e.g. specific conductance, pH, current velocity,

and dissolved oxygen) levels are certainly warranted.

Prairie chubs are endemic to the upper Red River basin, which is classified as a prairie stream system having, on average, higher specific conductance and lower dissolved oxygen than the lower Red River (Ruppel et al., 2017; Higgins & Wilde, 2005). Additionally, there are no physical barriers preventing the prairie chub from moving lower in the basin towards Lake Texoma in larger numbers, and conversely there is no physical barrier preventing shoal chubs from moving upstream (except for the Wichita River). Considering the potential importance of water quality in shaping aquatic fish communities (Barlow, 1958; Ostrand 2000; Ostrand & Wilde, 2001; 2004; Marie, Bernatchez, & Garant, 2012), this certainly presents an interesting hypothesis and warrants future experiments or studies to understand species restrictions.

Overall, I found an overall pattern of asymmetric hybridization, which could be due to the relative abundances of each species. In the zone of hybridization, shoal chubs are ~3x more abundant, thus providing more opportunities for reproduction with prairie chubs and hybrid individuals. This in turn has led to largely asymmetric introgression, with shoal chub alleles tending to introgress into individuals comprised of predominately prairie chub genomic backgrounds. This asymmetric introgression may be due in large part to many of the shoal chub alleles being selectively advantageous. This could be concerning from a conservation standpoint with respect to the genetic integrity of the pure prairie chub populations. However, in other riverine fishes, pairs of species with two independent hybrid zones have had different asymmetries with regards to introgressing alleles, which has been attributed to extrinsic or localized environmental selection pressures (Nolte, Gompert, & Buerkle, 2009; Aboim et al., 2010), which is not the case

here. Lastly, I found that assignment probability was predicted by various water quality parameters, indicating that the location of these species and their hybrids is highly associated with water quality. Thus, if the water quality were to change, potentially due to anthropogenic causes, which has been proposed in the past, such as attempting to decrease the salinity levels for agriculture use (U.S. Army Corps of Engineers 2012), I may see shifts in the species distributions which could be detrimental for the imperiled prairie chub. Finally, this study not only confirms that the prairie chub is a distinct lineage, supporting the nominal taxonomy (Eisenhour, 2004, Echelle et al., 2018; Underwood et al., 2003), it is also the first study to reveal extensive introgression between the shoal chub and prairie chub in the Red River basin of Texas which is associated with various environmental variables, predominately specific conductance.

Table 1.1. Collection locales of shoal and prairie chubs with sampling codes, river system, road crossing, latitude and longitude, sample sizes (N), and nucleotide diversity (π).

Code	River	Road Crossing	Latitude	Longitude	N	π
BR_1462	Brazos River	1462	29.34994	-95.58269	29	0.0038
BR_290	Brazos River	290	30.12944	-96.18693	30	0.0039
PR_104	Pease River	104	34.22786	-100.07375	8	0.0049
PR_283	Pease River	283	34.17915	-99.27841	7	0.0050
PR_287	Pease River	287	34.17982	-99.32343	19	0.0051
PR_6	Pease River	6	34.09471	-99.73016	27	0.0050
RR_259	Red River	259	33.68678	-94.69449	18	0.0053
RR_283	Red River	283	34.43119	-99.34181	17	0.0051
RR_35	Red River	35	33.72738	-97.15930	10	0.0057
RR_70	Red River	70	34.20985	-99.08233	10	0.0054
RR_79	Red River	79	34.13253	-98.09267	14	0.0057
RR_81	Red River	81	33.87807	-97.93435	14	0.0056
RR_89	Red River	89	33.91691	-97.51055	70	0.0055
WR_1919	Wichita River	1919	33.70029	-99.38871	66	0.0049
WR_6	Wichita River	6	33.82076	-99.78663	29	0.0049

Table 1.2. Pairwise genetic differentiation (Nei's G_{ST}) between all sites where genetic analyses were performed.

	BR_290	PR_104	PR_6	PR_283	PR_287	RR_89	RR_81	RR_35	RR_79	RR_70	RR_283	RR_259	WR_1919	WR_6
BR_1462	0.004	0.391	0.383	0.376	0.383	0.229	0.192	0.229	0.293	0.317	0.387	0.189	0.383	0.386
BR_290		0.385	0.376	0.370	0.377	0.224	0.187	0.224	0.287	0.311	0.380	0.184	0.377	0.379
PR_104			0.015	0.017	0.016	0.075	0.110	0.097	0.034	0.024	0.025	0.166	0.013	0.015
PR_6				0.009	0.008	0.068	0.103	0.090	0.027	0.016	0.017	0.158	0.005	0.006
PR_283					0.010	0.064	0.098	0.085	0.026	0.017	0.018	0.152	0.007	0.009
PR_287						0.069	0.103	0.090	0.028	0.017	0.018	0.159	0.006	0.008
RR_89							0.010	0.010	0.022	0.047	0.073	0.029	0.068	0.070
RR_81								0.014	0.043	0.073	0.107	0.020	0.103	0.105
RR_35									0.036	0.065	0.094	0.029	0.089	0.092
RR_79										0.020	0.034	0.081	0.026	0.028
RR_70											0.025	0.121	0.015	0.017
RR_283												0.162	0.015	0.017
RR_259													0.159	0.162
WR_1919														0.005

Table 1.3. Output from multiple regression of environmental variables (specific conductance, water depth, river kilometer, dissolved oxygen, current velocity, and pH) predicting q from Entropy (assignment probability to first genetic cluster from Entropy), including estimates slopes, t-values, p-values, percent R^2 explained, and upper and lower 95% confidence intervals for percent R^2 explained.

Term	Estimate	T-value	P-value	% R^2 Explained	Lower 95% CI	Upper 95% CI
Specific Conductivity	0.673	13.944	< 0.001	34.12	29.12	39.00
Depth	-0.374	0.092	< 0.001	29.71	26.32	32.71
River Kilometer	-0.099	-1.987	0.048	13.80	11.13	16.76
Dissolved Oxygen	-2.977	-10.020	< 0.001	17.74	12.48	23.15
Current Velocity	1.284	7.578	< 0.001	3.74	2.77	5.29
pH	6.496	4.056	< 0.001	0.89	0.28	3.05

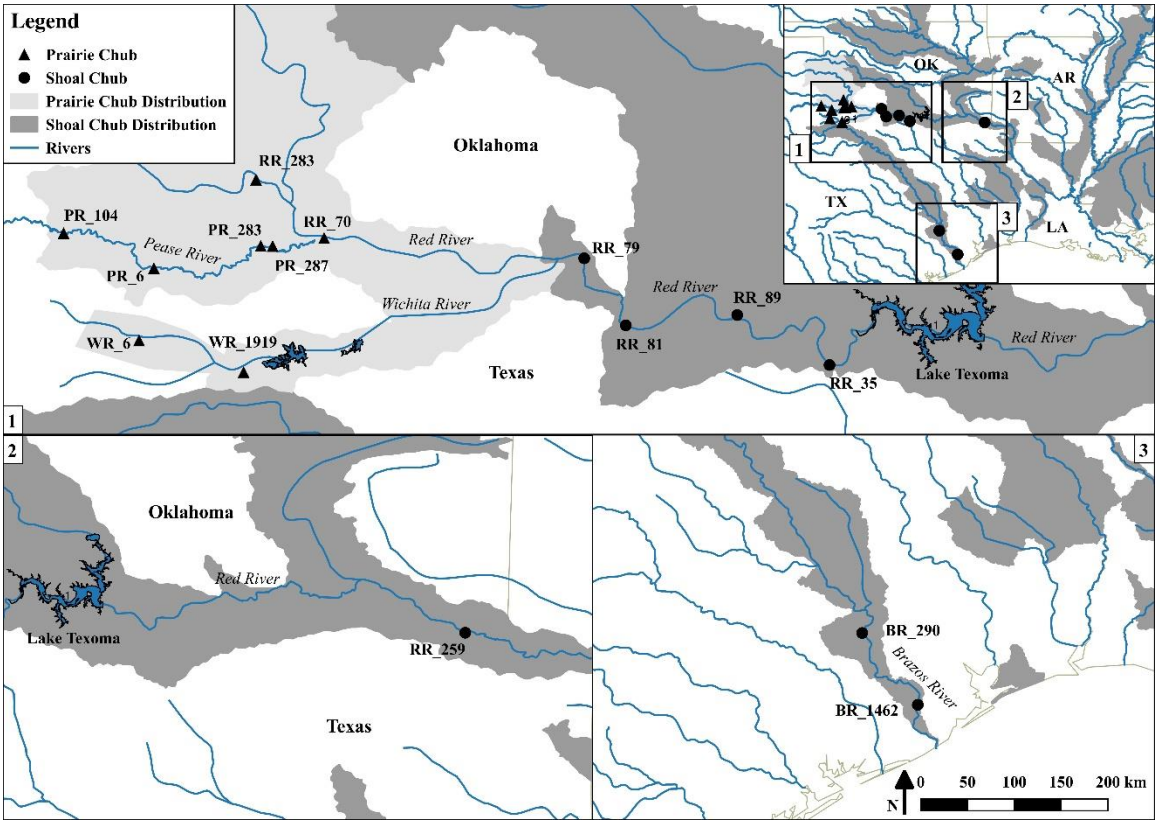


Figure 1.1. Map of locales where shoal chubs and prairie chubs were collected on the Red River (RR), Pease River (PR), Wichita River (WR), and Brazos River (BR). The map inset at the top right depicts the broad sampling frame statewide. Shapes denoting the sampling locations represent purportedly shoal chubs (circles) and prairie chubs (triangles) based on meristic morphological assignments. Light grey shading represents the prairie chub distribution, dark grey shading represents the shoal chub distribution (Data provided by NatureServe). Site codes are defined in Table 1.1.

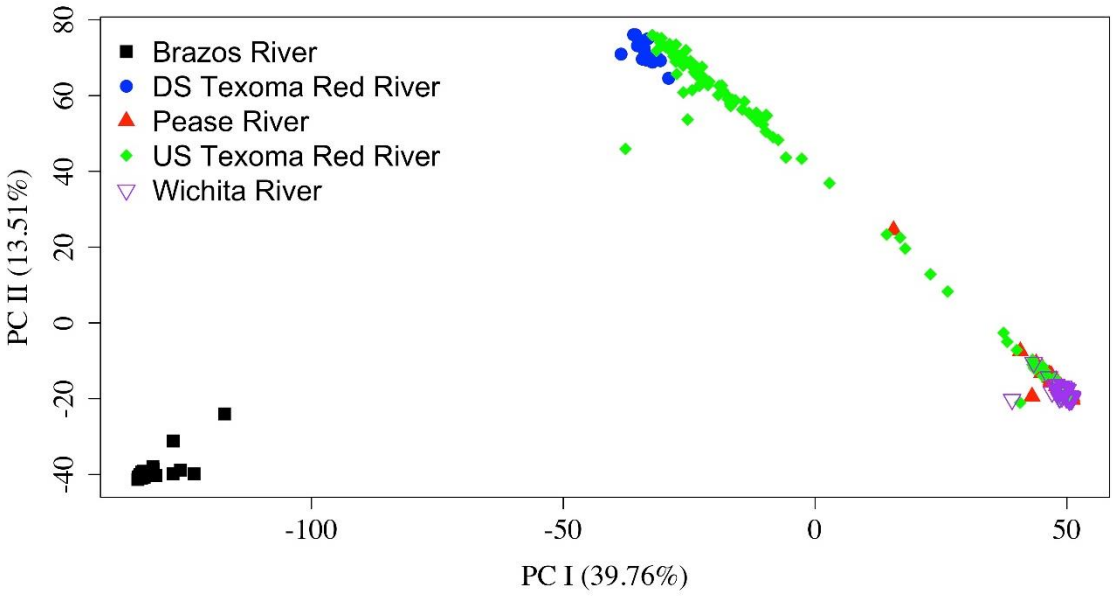


Figure 1.2. PCA of genetic differentiation of all individuals from each collection locale. PC I explains approximately 39.76% of the genetic variation in the data, with PC II explaining approximately 13.51% of the variation. Shapes and colors represent different collection rivers. Black squares are individuals from the Brazos River, blue circles are individuals captured downstream of Lake Texoma on the Red River, red triangles are individuals from the Pease River, green diamonds are individuals captured upstream of Lake Texoma from the Red River, and purple triangles are individuals captured from the Wichita River.

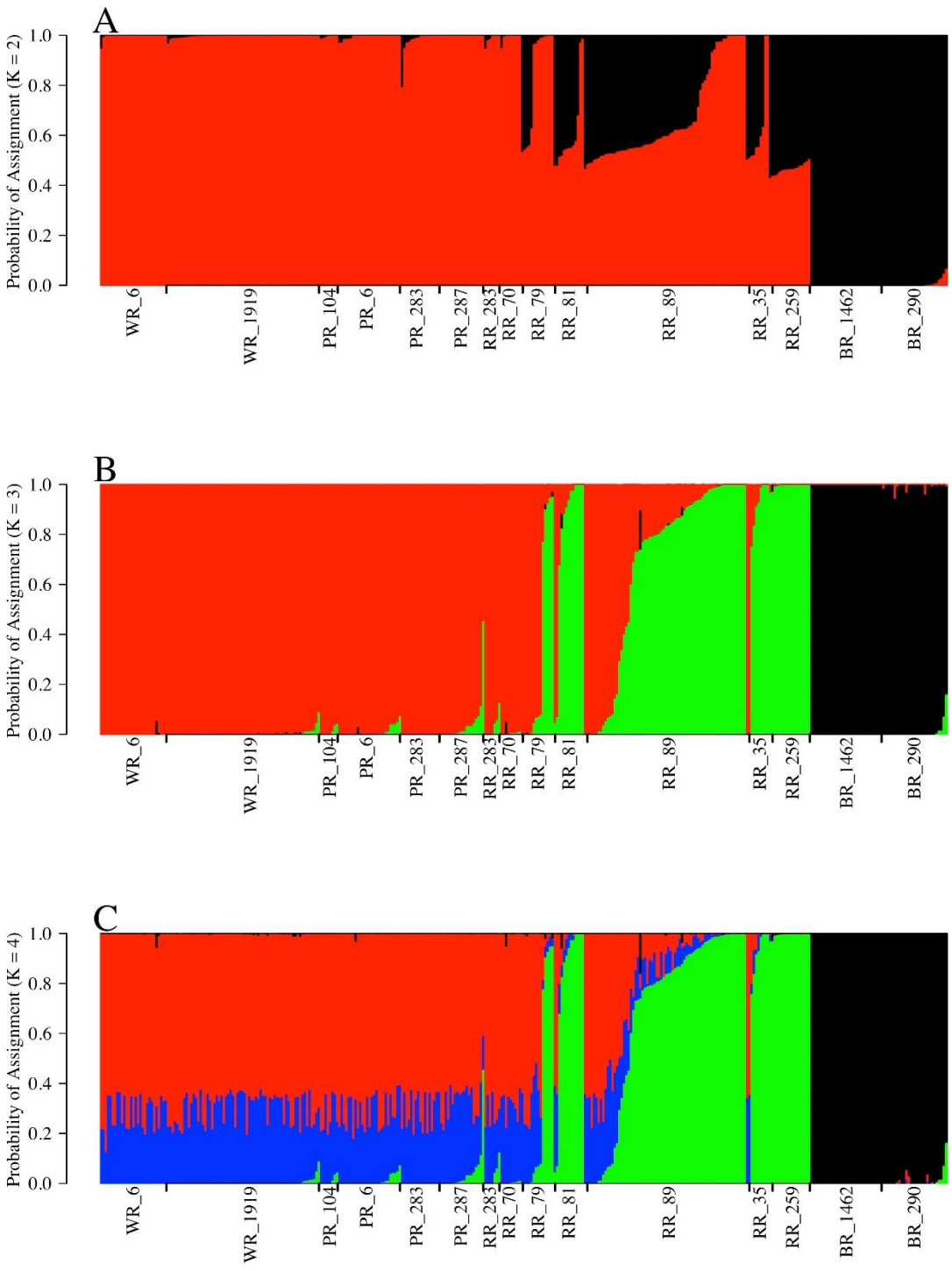


Figure 1.3. Entropy plots for k = 2 (A), 3 (B), and 4 (C) for all sampled sites oriented from upstream to downstream (left to right) in the Red River basin, followed by the Brazos River. Site codes are defined in Table 1.1.

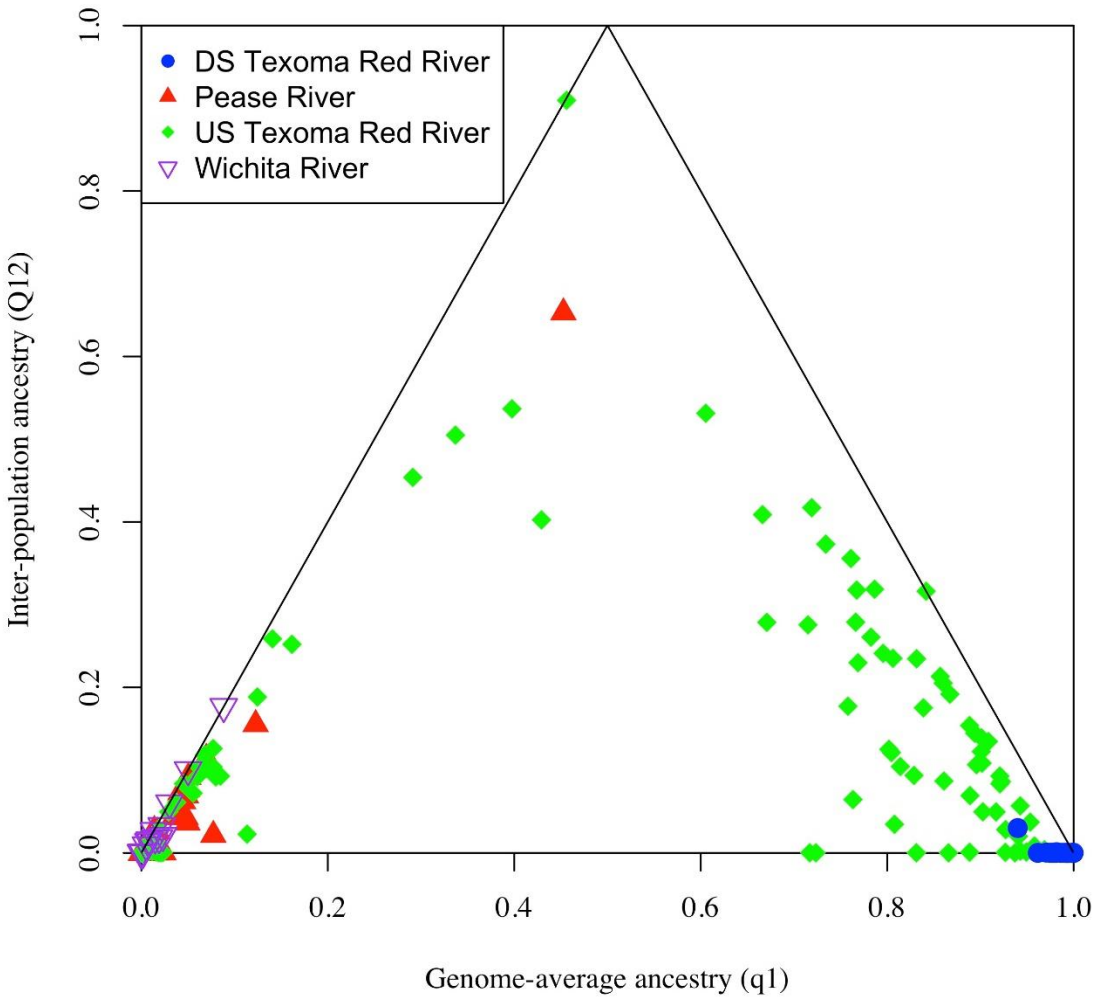


Figure 1.4. Scatter plot showing the relationship between genome-average ancestry (q1) and inter-population ancestry (Q_{12}). Symbols correspond to individuals from areas of the Red River basin (DS Texoma Red River = RR_259; Pease River = PR_287, PR_283, PR_6, and PR_104; US Texoma Red River = RR_35, RR_89, RR_81, RR_79, RR_70 and RR_283; Wichita River = WR_1919 and WR_6). Solid lines indicate the maximum possible inter-population population ancestry given global genetic ancestry.

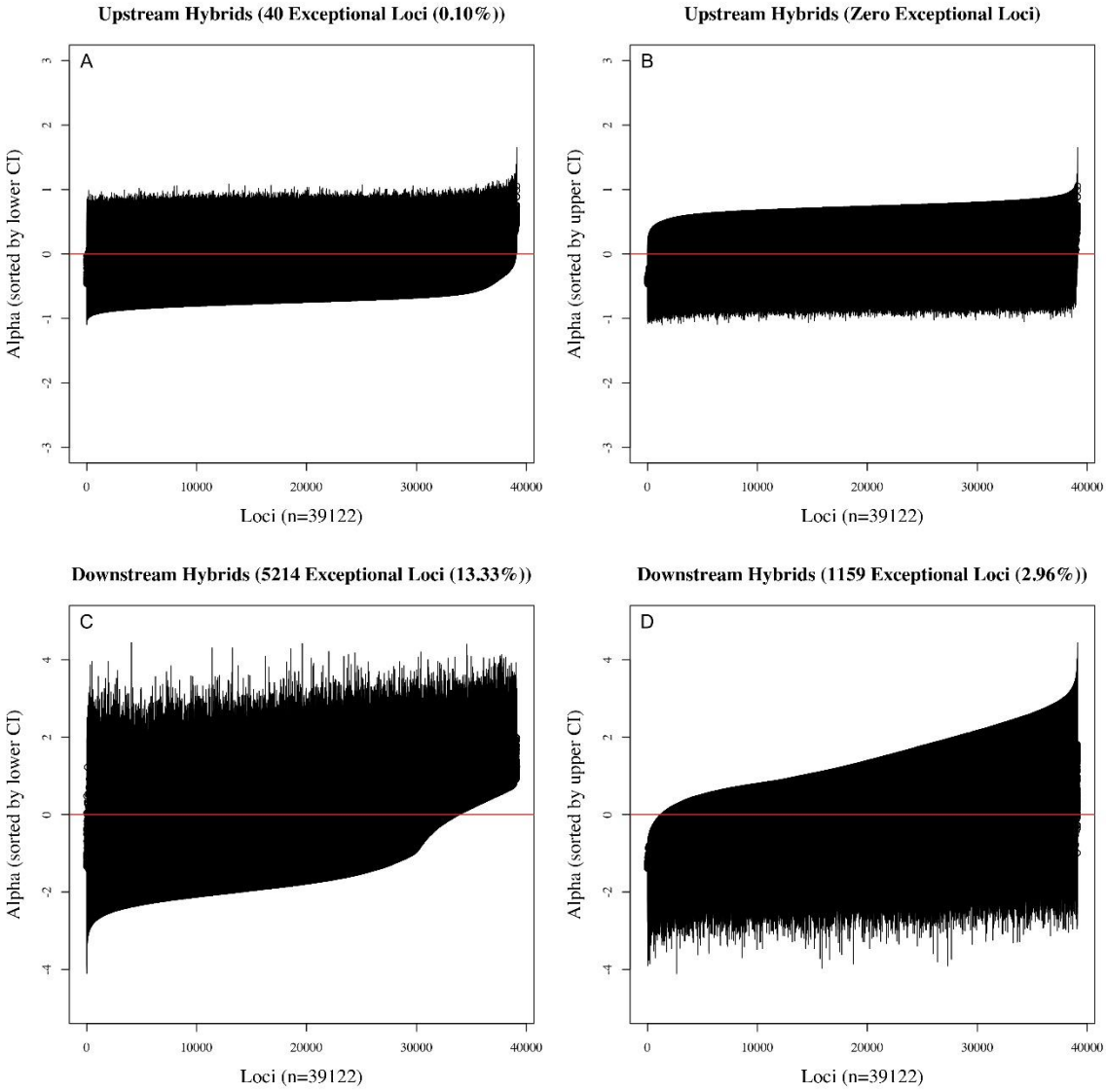


Figure 1.5. Median (± 95% CIs) of BGC cline parameters α sorted by the lower 95% CI and by the upper 95% CI for hybrid individuals captured in the upstream (A, B) and downstream Red River (C, D). Upstream hybrids include individuals captured at RR_283, PR_287, PR_6, and WR_1919; downstream hybrids include individuals captured at RR_89, RR_81, RR_35, and RR_79. Number of loci considered exceptional due to their 95% CI not overlapping zero are given in the title of each respective plot.

II. BIOGEOGRAPHY OF NORTH AMERICAN WESTERN GULF SLOPE FISHES: GENETIC RELATIONSHIPS REVEAL PATTERNS

Abstract

Understanding the patterns and processes which have led to present-day distributions of species is an important component of evolutionary biology and ecology. This study is a phylogeographic investigation of the *Macrhybopsis* complex (*M. hyostoma*, *M. australis*, *M. tetranema*, *M. marconis*, and *M. aestivalis*) throughout western Gulf of Mexico Slope drainages. A total of 24 sites on 11 rivers were sampled, which included river drainages from the Rio Grande to the Mississippi River. Genome-wide genetic data were utilized to assess their underlying genetic relationships and to interpret their genetic structuring in a historical biogeography context by testing specific biogeographical hypotheses. Population level and phylogenetic analyses revealed complex structuring, with eight total evolutionary lineages, three of which were undiscovered prior to this study. Additionally, there appear to be three major levels and one minor level of genetic structuring, corresponding with geography; the more distant and isolated lineages exhibiting higher divergence relative to lineages which are geographically closer. The majority of recognized species in this complex are restricted to a single basin, and the species which is widely distributed appears to have more complex and cryptic genetic structuring than previously recognized. Population level analyses and approximate Bayesian computation modelling support a stepping-stone model of dispersal, where coastal rivers which are more distant, isolated, and/or smaller relative to the source area of a species will be more genetically differentiated. This work supports the hypothesis that river drainages can be thought of as island-like systems with

increased lineage divergence within a single drainage.

Introduction

Phylogeographic investigations are important for understanding the patterns and processes which have led to present-day distributions of species (Berendzen, Simons, & Wood, 2003; Duvernell, Westhafer, & Schaefer, 2019; Echelle et al., 2005; Nagle & Simons, 2012; Osborne, Diver, Hoagstrom, & Turner, 2016; Walker, Stockman, Marek, & Bond, 2009). Assessing patterns and processes by testing models of dispersal can provide evidence for or against specific biogeographical hypotheses (Bennett et al., 2018; Boavida et al., 2019; Duvernell et al., 2019; Stobie, Oosthuizen, Cunningham, & Bloomer, 2018). Further, using genome-wide genetic data allows for fine-scale assessment of genetic structure, facilitating our understanding of biogeography in a robust and quantitative manner. Considerable work has been done on this topic in a limited number of geographic areas and scales (e.g. regional to continental), certain taxonomic scales (e.g. within and between species), terrestrial organisms, and aquatic organisms (Duvernell et al., 2019; Hughes et al., 2020; Machado, Galetti, & Carnaval, 2018; Nagle & Simons, 2012; Osborne et al., 2016; Walker et al., 2009), yet questions still remain with regards to specific modes of dispersal and evolution in fishes, especially those of western Gulf Slope drainages. This study uses genome-wide genetic data with population, phylogenetic, and model-based hypothesis testing using approximate Bayesian computation to explore the evolutionary history of a group of fishes in western Gulf Slope drainages.

Fish diversity is highest in the Southeastern USA in coastal drainages that were connected during sea level fluctuations and left largely unglaciated during the Pliocene

and Pleistocene (Conner & Suttkus, 1986; Jenkins, Van Houtan, Pimm, & Sexton, 2015). The costal drainages of the western Gulf Slope are vastly different today than during the Pliocene and Pleistocene: some historically large rivers no longer exist, some rivers were much larger, while other rivers had different flow patterns and pathways (Galloway, Whiteaker, & Ganey-Curry, 2011). There were periods of lowered sea level which created coastal connections for many rivers, and during glacial maxima the southern USA was wetter allowing for further river connections or river captures (Conner & Suttkus, 1986; Galloway et al., 2011; Mayden, 1985, 1988). These historical connections and river pathways led to dispersal between basins for fishes followed by increased rates of vicariance when the connections were severed (Conner & Suttkus, 1986; Mayden, 1985, 1988; Sepkoski & Rex, 1974). However, the specific model of dispersal remains unclear for many species' complexes within Gulf Slope drainages.

A general model of fish diversification for dispersal between drainages would predict that drainages increasingly distant to the epicenter of origin of fish clades would exhibit increasing genetic divergence (Eckert, Samis, & Lougheed, 2008; Le Corre & Kremer, 1998; Sepkoski & Rex, 1974). This prediction would be supported by a stepping-stone model of dispersal (Le Corre & Kremer, 1998; Sepkoski & Rex, 1974) and the central-marginal hypothesis (Brussard, 1984; Eckert et al., 2008). Both the stepping-stone model of dispersal and the central-marginal hypothesis posit that species or populations in areas (e.g. rivers for fishes) which are further from the epicenter of a species will be more genetically distinct (Eckert et al., 2008; Le Corre & Kremer, 1998); however, there are studies which refute these ideas (Eckert et al., 2008). In this study, the stepping-stone model of dispersal is treated as such, a model of dispersal, whereas the

central-marginal hypothesis can be a consequence of this and other types of dispersal. Additionally, the movement of fishes between drainages can be treated as an island-like system, in that fishes cannot move between drainages in the absence of a connection (Carvajal-Quintero et al., 2019; Dias et al., 2014; Sepkoski & Rex, 1974). Thus, historical river connectivity can be a primary determinant of current species ranges (Carvajal-Quintero et al., 2019; Dias et al., 2014) and species that inhabit drainages connected during the lowered sea levels of the Quaternary have comparatively larger range sizes than those species which do not (Carvajal-Quintero et al., 2019).

In addition to historical river connectivity leading to the historical diversification and current distribution of fishes, water permanency can also play a role, acting as evolutionary refugia (Bernatchez & Wilson, 1998; Craig, Kollaus, Behen, & Bonner, 2016; Soltis, Morris, McLachlan, Manos, & Soltis, 2006). Southern areas of the USA were unglaciated during glacial maxima and organisms are hypothesized to have generally been able to persist through glacial cycles, unlike in more northern areas. Thus, there is a negative relationship between latitude and genetic differentiation within and between species (April, Hanner, Mayden, & Bernatchez, 2013). This relationship suggests that fishes in more southern latitudes have been persisting and diverging from other species longer than those in northern latitudes and provides a prediction that more southerly species will be more diverged. Persistence followed by subsequent divergence was likely occurring during the glacial minima when the Southern USA was much drier, as some fishes were able to persist in basins which remained wet, likely acting as evolutionary refugia (Bernatchez & Wilson, 1998; Soltis et al., 2006). These areas which have been proposed as being evolutionary refugia are potentially responsible for the high

levels of endemism in fishes (Jetz, Rahbek, & Colwell, 2004; Oberdorff, Lek, & Guegan, 1999; Tedesco et al., 2012), including those in western Gulf Slope drainages (Craig et al., 2016).

Research in western Gulf Slope drainages has revealed that historical river connectivity is likely responsible for the spread of many fishes. For example, the pupfish (*Cyprinodon spp.*; Echelle et al., 2005), red shiner (*Cyprinella lutrensis*; Osborne, Diver, Hoagstrom, & Turner, 2016), topminnows (*Fundulus notatus* complex; Duvernell, Westhafer, & Schaefer, 2019), river chubs (*Nocomis spp.*; Nagle & Simons, 2012), northern hogsuckers (*Hypentelium nigricans*; Berendzen, Simons, & Wood, 2003), and multiple species groups in the Ouachita Highland area (Mayden, 1985) have all reached their present distributions due to historical connectivity of river systems. Another group which presents a powerful system to test biogeographic hypotheses of historical connectivity is the *Macrhybopsis* species complex (Echelle et al., 2018; Eisenhour, 2004; Gilbert, Mayden, & Powers, 2017). *Macrhybopsis* likely evolved in turbid streams, as these fishes have barbels on the corners of their mouths that function as sensory organs used to find food (Eisenhour, 2004). Given that, it has been hypothesized that the genus had a Mississippi-Missouri River origin in turbid systems in glacial and preglacial Plains rivers (Eisenhour, 2004). This proposed origin is supported by the geographic distribution of members of the genus as the majority of species occur within the Mississippi River basin, and those that do not largely occur in coastal or adjacent drainages (Figure 2.1). The common ancestor of current *Macrhybopsis* species was hypothesized to have dispersed during wet periods throughout the Pliocene or Pleistocene and to have established in rivers throughout the United States (Eisenhour, 2004). Once this ancestor

was distributed across largely present-day drainages (e.g. Brazos, Colorado, Red, and Rio Grande river basins), populations potentially became isolated due to costal connections being severed or during glacial retreats during the Pleistocene. Due to the present-day distribution of *Macrhybopsis*, members of this genus might exhibit genetic structure that provides evidence for this stepping-stone model of dispersal (Le Corre & Kremer, 1998; Sepkoski & Rex, 1974) as well as the central-marginal hypothesis (Brussard, 1984; Eckert et al., 2008).

Recent work on this genus and other small riverine fishes has revealed substantial cryptic diversity, with high genetic divergence within species between basins (Duvernell et al., 2019; Echelle et al., 2018; Gilbert et al., 2017; Nagle & Simons, 2012; Osborne et al., 2016). The *Macrhybopsis* complex constitutes an ideal system in which to test biogeographical hypotheses using model-based approaches and to evaluate the possibility of cryptic divergence of fishes within Gulf of Mexico Slope drainages. Within the Gulf Slope drainages in Texas, there are five recognized *Macrhybopsis* species that generally occur in separate river basins. *Macrhybopsis hyostoma* occurs in the Red, Brazos, Clear Fork, and Colorado Rivers, however this species exhibits complex genetic structuring and may not represent a single evolutionary lineage (Echelle et al., 2018; Gilbert et al., 2017; Underwood et al., 2003). *Macrhybopsis tetranema* occurs in the Canadian River of the Arkansas River basin; *M. australis* occurs in the upper Red River and its tributaries, Pease and Wichita rivers, in the Mississippi River basin; *M. marconis* occurs in the Colorado, Guadalupe, and San Antonio river basins; and *M. aestivalis* occurs in the Rio Grande basin. Additionally, three of these species have restricted ranges within their respective basins, with *M. australis* restricted to the upper reaches and tributaries of the

Red River, *M. marconis* occurring within a portion of the Colorado River basin on the Edwards Plateau, and *M. hyostoma* occurring in the lower portion of the Colorado and Red rivers (Eisenhour 2004). *Macrhybopsis* thus present an interesting case study into how fishes might have dispersed and diverged between and within Gulf Slope drainages due to their wide and disparate distribution.

The goal of this study was to assess the genetic structure and phylogeographic relationships of *Macrhybopsis* species within and across western Gulf Slope drainages. In this study, genotyping-by-sequencing (GBS) techniques were used to generate 20,376 SNPs for population and lineage level analyses. The first objective of this study was to quantify the genetic structure of *Macrhybopsis* within and between western Gulf Slope drainages using population level analyses. The second objective of this study was to use approximate Bayesian computation (Do It Yourself Approximate Bayesian Computation, DIYABC) to test evolutionary divergence scenarios. The tested scenarios were created based off this and other published work, while also explicitly testing various biogeographical hypotheses (e.g. variations of peripheral isolation, vicariance, a stepping-stone mode of dispersal, and the central-marginal hypothesis). Third, the resolution of the data in this study and the results presented allow for a discussion of cryptic diversity, with implications for increasing our understanding of the biodiversity within small-bodied riverine fishes.

Methods

Fishes were collected using seines from 24 sites on 11 rivers (Table 2.1; Figure 2.1). *Macrhybopsis aestivalis* was collected from the Rio Grande, *M. australis* was collected from the Red, Pease, and Wichita rivers, *M. hyostoma* was collected from the

Colorado, Brazos, Clear Fork Brazos, Red, and Mississippi rivers, *M. marconis* was collected from the Colorado, Guadalupe, and San Antonio rivers, and *M. tetranema* was collected from the Canadian River. Specimens were euthanized with Tricaine Methanesulfonate (MS-222, Western Chemical, Inc.) and then immediately preserved in 95% ethanol to minimize DNA degradation.

Genomic DNA was extracted from preserved fin clips in 96-well format using Qiagen DNEasy blood and tissue extraction kits and prepared for genotyping. Reduced representation libraries were prepared following genotype-by-sequencing protocols following Gompert et al., 2014, Parchman et al., 2012, and Sotola, Ruppel, Bonner, Nice, & Martin, 2019. Briefly, DNA from each individual was digested with the restriction enzymes EcoR1 and Mse1 (New England Biolabs; NEB, Inc.) and were labeled by ligating Illumina adapters with unique 8-10 base pair sequences onto the fragmented DNA. Two separate rounds of PCR were performed on these restriction-ligation products using Illumina primers, and the final PCR products were pooled into a single multiplexed library. This library was then sent to the University of Texas Genomic Sequencing and Analysis Facility (Austin, TX, USA) and sequenced over two lanes on an Illumina HiSeq 2500 SR 1x 150 platform after size selection for fragments between 300 and 400 base pairs using a Pippen Prep quantitative electrophoresis unit (Sage Science, Beverly, MA). Two libraries were prepared, sequenced, and combined for this dataset.

After sequencing, resulting raw reads were processed first by identifying and removing PhiX control sequences using Bowtie version 1.1.2, then assembled to the PhiX genome (Langmead & Salzberg, 2012). Using a custom Perl script reads from Mse1 adapters and barcodes were removed single-base sequencing mutations were corrected in

barcodes and sample IDs were matched with unique barcode identifiers. Because a reference genome is not available for *Macrhybopsis*, a *de novo* assembly was performed using part of the dDocent variant calling pipeline (Puritz, Hollenbeck, & Gold, 2014). Specifically, unique reads were found for each individual, and reads with less than four copies and shared across less than four individuals were filtered out of the dataset. The resulting filtered reads were assembled using CD-hit (Fu, Niu, Zhu, Wu, & Li, 2012) with a threshold of 92% similarity. The consensus reads from this *de novo* assembly formed the basis for a reference-based assembly. All sequence reads were assembled to the reference using Burrows Wheeler Aligner (BWA 0.7.13-r1126; Li et al., 2009) allowing up to four mismatches. BCFtools version 1.9 was used to identify variable sites (Single Nucleotide Polymorphisms – SNPs) using the *mpileup* and *call* commands, ignoring indels and only calling biallelic loci; likelihoods for genotypes for every individual and for every variable site were calculated (Li et al., 2009). Filtering of variable sites was performed using custom Perl scripts with filtering thresholds to exclude variable sites set as follows: sequence depth less than 2X coverage and greater than 15,386 reads (equal to mean sequence depth across sites plus two times the standard deviation to filter out potential paralogs), mapping quality less than 30, an absolute value of the mapping quality rank sum test greater than 2.5, absolute value of the read position rank sum test greater than 2, absolute value of the base quality rank sum test greater than 3, minor allele frequency less than 0.05, and missing data for more than 50% of individuals. Subsequently, for contigs containing more than one SNP, only a single randomly chosen SNP was selected and included in the final data matrix to minimize linkage disequilibrium.

The number of sampled individuals was skewed heavily towards the Red River basin due to sampling for a previous study on hybridization between *M. hyostoma* and *M. australis* (Sotola et al., 2019). Due to this, initial sample sizes for the various putative lineages varied substantially (e.g., we had a total of 241 *M. australis*, whereas we only sampled a total of 21 *M. hyostoma* from the Mississippi River and 13 from the Clear Fork Brazos River). Thus, due to potential issues with unbalanced sample sizes, downsampling of individuals was performed (Haselhorst, Parchman, & Buerkle, 2019; Meirmans 2019). After preliminary Entropy analyses revealed the general genetic structure patterns in these samples, each identified lineage was downsampled to approximately the same number of individuals (Table 2.1). Additionally, hybrid individuals were filtered out, as those data were not within the scope of this study (Sotola et al., 2019). After filtering of individuals with low median coverage, hybrid individuals, and downsampling, a total of 189 individuals were used in downstream analyses.

To quantify genetic structure of the *Macrhybopsis* complex, admixture proportions for each individual, allele frequencies for each population, and posterior genotype probabilities were estimated using Entropy (Gompert et al., 2014). Entropy is a Bayesian clustering algorithm that accounts for uncertainty in sequence coverage, sequence alignment, and genotyping errors by incorporating maximum likelihood estimates of genotypes. Models with different numbers of populations (k = 2 – 10) were run, and no attempts were made to identify the "best" k because examination of all levels of k provide a more comprehensive understanding of population structure (Driscoe et al., 2019; Gilbert et al. 2012; Janes et al. 2017; Meirmans, 2015; Sotola et al., 2019). Posterior distributions of genotypes and admixture proportions were calculated for each k

using Markov Chain Monte Carlo (MCMC) with 200,000 iterations. The first 5,000 iterations were discarded and then sampled every 10^{th} iteration, and each model for all k clusters were run two times (i.e. two MCMC chains). The Gelman-Rubin diagnostic statistic and effective sample sizes were calculated to verify chain convergence, and posterior genotype estimates and admixture proportions were averaged across both chains. Posterior distributions for parameters were summarized as means, medians, and 95% credible intervals (CIs).

To explore population differentiation among all clusters for k=8, pairwise Nei's G_{ST} (Nei, 1987) was calculated among all lineages. Allele frequencies were calculated in R (R Core Team, 2019) from the mean genotype posterior probabilities, which were in turn used to calculate pairwise G_{ST} values. In addition, population-level variation for each lineage was reported using the genetic diversity index (π) and the number of segregating sites (θ) calculated with SAMtools using the expectation-maximization (EM) algorithm, employing 20 iterations for each collection locale to achieve convergence of estimates (Li, 2011). Resulting π and θ estimates were plotted as a barplot. In order to summarize the distribution of genetic variation, Principal Component Analysis (PCA) was performed in R (R Core Team, 2019) on the genetic covariance matrix calculated from the genotype probability estimates generated in Entropy (Gompert et al., 2014).

To assess phylogenetic relationships of all lineages identified in Entropy, SNAPP (Bryant, Bouckaert, Felsenstein, Rosenberg, & Roychoudhury, 2012) was implemented in BEAST (Drummond & Rambaut, 2007). An initial run was performed to confirm the existence of the eight lineages supported by Entropy (see results) by randomly selecting two individuals from each lineage (16 total individuals), then running the analysis with

each individual as its own taxon (i.e. 16 total taxa). Results from this analysis confirmed the presence of the eight lineages, so individuals from each lineage were combined in the final analysis. Due to the intense computational time of this analysis, individuals and SNPs were filtered further. Each of the eight lineages were subsampled to three individuals, which was based off membership in each of the groups of the k=8 Entropy plot and preliminary SNAPP analysis: *M. aestivalis*, *M. australis*, Brazos Colorado *M. hyostoma*, Clear Fork *M. hyostoma*, Red *M. hyostoma*, Mississippi *M. hyostoma*, *M. marconis*, and *M. tetranema*. Further, SNPs were filtered with the following thresholds: minimum coverage at 20X, and only one individual was allowed to have missing data at a variable site, all other filtering thresholds were set as in the full dataset. The single Markov chain Monte Carlo (MCMC) chain length was 10,000,000, with sampling every 1,000 iterations, with no burn-in, and using default priors and model parameters provided from the program. Tracer (Drummond & Rambaut, 2007) was used to assess mixing with adequate effective sample sizes (> 200) for all parameters. Phylogenetic tree was visualized with DensiTree (Bouckaert, 2010).

To assess the probability of several evolutionary divergence scenarios among lineages, DIYABC was implemented (Cornuet et al., 2014). DIYABC uses approximate Bayesian computation methods to simulate large datasets under a given set of evolutionary scenarios, from which the most probable divergence scenario can be determined. Population genetic summary statistics (e.g. proportion of zero or mean of non-zero values of F_{ST} and Nei's D) are calculated from the simulated datasets and compared to the observed dataset which allows for inference and comparison of alternative evolutionary scenarios. The full SNP dataset (N = 20,376 SNPs) was

subsampled prior to running this analysis to reduce computational time as follows. To determine the minimum number of SNPs required which still captured the standing genetic variation within the full dataset, random datasets were drawn from the full dataset. Random datasets consisted of draws of 50, 100, 250, 500, and 1000 SNPs, with 10 random datasets total for each number of SNPs (e.g. 10 random draw datasets of 50 SNPs). PCAs were performed on each random dataset separately, then were compared using a Procrustes analysis to determine the degree of similarity between PCA plots of subsampled SNPs and the full dataset. The highest degree of similarity with the lowest number of randomly sampled SNPs of the curve was found at 250 SNPs, however 500 SNPs were used to ensure that genetic structure was still captured in this subsampled dataset while still allowing reasonable computation time (Table 2.2). The Procrustes analysis was performed with the *vegan* packages in R (Oksanen et al. 2019).

The program DIYABC can only use "called" SNPs (as opposed to genotype likelihoods, which are used for Entropy, or genotype probabilities which are produced from Entropy), thus genotype probabilities were rounded and called. Genotype probabilities are scaled from 0 to 2, so thresholds of ≥ 0.0 & ≤ 0.66 were rounded to 0, ≥ 0.67 & ≤ 1.33 were rounded to 1, and ≥ 1.34 were rounded to 2. This threshold and others were compared to the full dataset of genotype probabilities via Procrustes analysis, and the above threshold was the most similar (Procrustes R = 0.997). Additionally, to further reduce computational time, 10 individuals from each of the eight lineages were randomly selected. An additional Procrustes correlation was run on the PCA of the final dataset of 500 randomly selected SNPs versus the total dataset, and was found to be highly similar (Procrustes R = 0.984). In total, 10,000,000 simulations were run to generate datasets to

compare divergence scenarios.

A total of 16 scenarios were tested (Table 2.3; Appendix 2.1). Scenarios were designed based on hypotheses derived from several sources, including results from this study (e.g. Entropy, G_{ST} relationships, and SNAPP; see results), previous publications on *Macrhybopsis* (Echelle et al., 2018; Eisenhour, 2004; Underwood et al., 2003), and various models of dispersal and vicariance. Specifically, the set of scenarios includes a null model of simultaneous divergence (i.e. basal polytomy for all lineages), and several scenarios were designed to match the topology of the phylogenetic analysis, Entropy plots, and G_{ST} relationships that comport with results from the present study. Additional scenarios were created to match the topology of phylogenetic analyses from Echelle et al., 2018 and Underwood et al., 2003. Finally, several scenarios were designed based off of general patterns of vicariance and a stepping-stone model of dispersal (Eisenhour, 2004; Sepkoski & Rex, 1974). The posterior probabilities of the various scenarios were compared using a direct approach, which compares the relative proportion (top 1%) of each scenario in the simulated data sets that are closest to the observed data, and a logistic regression, which was run on the deviations between simulated and observed summary statistics (Beaumont, 2008; J. M. Cornuet et al., 2008; Fagundes et al., 2007).

Results

A total of 776 initial individuals were captured, 384 of which were sampled and included in a previous study (Sotola et al., 2019). After SNP filtering, removal of low coverage individuals with a mean of < 2 reads per locus, and downsampling of localities used in previous studies, a total of 189 individuals and 20,376 SNPs were analyzed in the current population level analyses (e.g. G_{ST}, PCA, Entropy; Table 2.1). For the

phylogenetic analysis, the dataset used in SNAPP consisted of 24 individuals, three from each lineage, and included a total of 3,134 SNPs. To test biogeographical hypotheses using DIYABC, 80 individuals, 10 from each lineage were included with a total of 500 SNPs. See Table 2.1 for lineage abbreviations.

Overall, there was a large amount of genetic differentiation and structuring between all pre-identified and newly identified lineages within the *Macrhybopsis* complex. The highest G_{ST} values were found between *M. marconis* and *M. aestivalis* and all other lineages (range from 0.42 to 0.52; Figure 2.2). G_{ST} values between the BrCol *M. hyostoma* and other lineages ranged from 0.26 to 0.35. The lowest was between C.F. *M. hyostoma* and BrCol *M. hyostoma* at 0.03, and M.R. *M. hyostoma* and R.R. *M. hyostoma* 0.05. *M. australis* had moderate pairwise levels of G_{ST} with *M. tetranema* (0.11), M.R. *M. hyostoma* (0.015), and R.R. *M. hyostoma* (0.11). Highest π were found in the M.R. (0.009), C.F. (0.008), and R.R. (0.007) *M. hyostoma*, followed by *M. australis* (0.006) and *M. tetranema* (0.006); lowest was found in *M. marconis* (0.003; Figure 2.3). Highest θ was found in M.R. (0.013) and C.F. (0.012) *M. hyostoma*, *M. aestivalis* (0.012), and *M. marconis* (0.012); lowest was in the BrCol *M. hyostoma* (0.007).

PC axes one through 4 explained 77.8% of the variation in the SNP matrix and plots reveal substantial structuring within this complex. PC axis one explains 33.66% of the variation, splitting *M. aestivalis* and *M. marconis* from all other lineages (Figure 2.4A), while PC axis two explains 21.51% of the variation and splits BrCol and C.F. *M. hyostoma* from M.R. and R.R. *M. hyostoma*, *M. australis*, and *M. tetranema*. PC axis three explains 17.11% of the variation and distinguishes *M. aestivalis* from *M. marconis* (Figure 2.4B). Finally, PC axis four explains 5.52% of the variation in the data and splits

out the remaining lineages (Figure 2.4C).

Entropy plots reveal similar clustering to the PCA plots (Figure 2.5). The first split at k=2 groups *M. aestivalis* and *M. marconis* together separately from all other lineages. At k=3, BrCol and C.F. *M. hyostoma* rivers split. At k=4, *M. marconis* splits from *M. aestivalis*. At k=5, *M. tetranema* and *M. australis* split from the M.R. and R.R. *M. hyostoma*. At k=6, there is a group which appears in the BrCol and C.F. *M. hyostoma*. This cluster appears predominately in individuals from the Brazos River drainage, showing very low proportions (< 0.05) in individuals from the Colorado River. At k=7, *M. australis* splits from *M. tetranema*, and at K=8 R.R. *M. hyostoma* splits, mostly, from M.R. *M. hyostoma*. At k=9 the model appears to break down as no biologically interpretable cluster is added.

Mean admixture proportions from Entropy for each lineage were plotted as pie charts and placed on a map to explore the splits and patterns from a biogeographical perspective (Figure 2.6). Geographically, at k=2 the split results in a northeastern and southwestern group, with the two lineages which are more geographically distant split out together from the other lineages. At k=3, the BrCol and C.F. *M. hyostoma* split from the northern drainages. At k=4, the two isolated lineages split, *M. aestivalis* and *M. marconis*. At k=5, there is an east (*M. australis* and *M. tetranema*) and west (*M. hyostoma* Red and Mississippi) split within the Mississippi River drainage. At k=6, there is a slight split within the Brazos drainage, with the Clear Fork having majority of the new cluster. At k=7, the upper Red River basin, *M. australis*, splits from *M. tetranema* in the Canadian. Lastly, at k=8 *M. hyostoma* Red River splits from *M. hyostoma* Mississippi River.

The phylogenetic tree produced via SNAPP was largely in concordance with the

other analyses (Figure 2.7). *Macrhybopsis marconis* and *M. aestivalis* are the most diverged relative to the other lineages. Subsequently, B.R. and C.F. *M. hyostoma* split from the Mississippi River basin (e.g. *M. australis*, *M. tetranema*, M.R. and R.R. *M. hyostoma*). Among lineages within the Mississippi River basin, M.R. *M. hyostoma* splits first, followed my *M. tetranema*, then R.R. hyostoma and *M. australis*.

After running ABC simulations to compare evolutionary divergence scenarios, each of the comparison analyses (i.e. direct and logistic approach) indicated different scenarios as having the highest probability. The scenario which was most likely as indicated by the direct approach was based on results of the Entropy analysis (Figure 2.8A, Appendix 2.2A). In contrast, the scenario which was most likely as indicated by the logistic approach was based off a stepping-stone model of dispersal and vicariance (Figure 2.8B, Appendix 2.2B).

Discussion

Within the *Macrhybopsis* complex, currently consisting of five recognized species, there appears to be genetic structuring at a variety of scales. Within the lineages uncovered in this study, the highest and first level of structure is between the distant and isolated lineages relative to the epicenter of the complex (e.g. *M. aestivalis* and *M. marconis*, compared to all other lineages) supporting the central-marginal hypothesis (Eckert et al., 2008). The next level of genetic structure is between lineages that are both geographically closer to one another and closer to the putative epicenter of the group's origin, although still in separate coastal basins (e.g. Brazos and Colorado *M. hyostoma* with Mississippi River lineages). The last level of structure is within the Mississippi River drainage, the presumed epicenter and origin (e.g. *M. australis*, *M. tetranema*, and

Red and Mississippi *M. hyostoma*). The majority (e.g. *M. australis*, *M. tetranema*, and *M. aestivalis*) of recognized species in this complex are restricted to a single basin, and the species that is widely distributed (e.g. *M. hyostoma*), has more complex and deeper genetic structuring than previously recognized. Levels of genetic diversity are similar among lineages, however, the two lineages which are the most geographically isolated, *M. aestivalis* and *M. marconis*, have the lowest levels of nucleotide diversity. Across all analyses, our results suggest these eight lineages have an evolutionary history that involves complex patterns of dispersal and vicariance within Gulf Slope drainages which potentially acted as evolutionary refugia during glacial cycles. Additionally, this work suggests that much more cryptic diversity exists in riverine fishes, especially in Gulf Slope drainages, and adds to our growing knowledge of the evolutionary history of fishes.

In all analyses, the highest levels of structuring (i.e. deepest divergence) was found between *M. aestivalis*, *M. marconis* and all other lineages. *Macrhybopsis aestivalis* and *M. marconis* are the two most geographically distant and isolated relative to all other lineages (Figure 2.1). *Macrhybopsis aestivalis* is restricted to the Rio Grande basin, furthest southwest, whereas *M. marconis* is restricted to the Guadalupe and San Antonio rivers and parts of the Colorado River. Due to their distance from the Mississippi River basin epicenter of this genus and high levels of genetic divergence, it's likely these lineages have been isolated for a relatively longer period of time. The next deepest level of genetic structuring occurs within and between *M. hyostoma* lineages, *M. australis*, and *M. tetranema*. These lineages have moderate levels of genetic differentiation, indicating incomplete divergence, but each split out as an entity in the phylogenetic analysis, which

suggests more recent divergence (i.e. separation) relative to *M. aestivalis* and *M. marconis*. The moderate level of divergence is likely due to the fact that the drainages where these lineages occur are closer geographically and have been connected more recently (Conner & Suttkus, 1986). The *M. hyostoma* lineages from the Mississippi, Red, Brazos, and Colorado rivers were treated as one entity until recently (Echelle et al., 2018). However, recent work with mitochondrial and nuclear markers has raised doubts and found divergence between *M. hyostoma* in different basins, even outside of Gulf Slope drainages (Echelle et al., 2018; Gilbert et al., 2017), which the present results support. Thus, it is interesting to see similar levels of genetic differentiation between Brazos and Colorado *M. hyostoma* and Red and Mississippi *M. hyostoma* relative to their genetic differentiation levels with *M. australis* and *M. tetranema*.

Lineages within the Mississippi River basin (e.g. Red, Canadian, and Mississippi rivers) show lower levels of differentiation and less structuring. These lineages are currently in the same river basin; thus, it is unsurprising they show lower levels of genetic structuring. Although present day gene flow is unlikely due to barriers (e.g. dams), it is likely these lineages share a more recent ancestor which was able to disperse throughout the Mississippi River basin prior to divergence. This is supported by the present and other work, showing that *M. hyostoma* within a drainage is more genetically similar to the endemic species (e.g. *M. australis*) than other *M. hyostoma* (Echelle et al., 2018; Gilbert et al., 2017; Underwood et al., 2003). This pattern of relationships suggests that cryptic diversity is widespread in the *Macrhybopsis* genus, and by extension potentially in small-bodied riverine fishes in general, and could be prominent in other widespread minnows within Gulf Slope drainages.

The lowest levels of genetic differentiation and structuring are found within *M. hyostoma* in single rivers. Interestingly, Red *M. hyostoma* show very low differentiation and incomplete divergence from Mississippi *M. hyostoma*, even though the sampled populations are geographically distant, although still in the same basin. Clear Fork *M. hyostoma* had the lowest differentiation and incomplete divergence from Brazos and Colorado *M. hyostoma*. However, both lineages split as their own entity in the phylogenetic analysis. These lineages could be more recently diverged and undergoing similar processes of divergence to *M. australis* and *M. tetranema*, as there are similarities with these lineages in the Red and Canadian (Arkansas River basin) rivers which will be explored further below.

The population-level analyses and the coalescence divergence scenarios presented in this study support a stepping-stone mechanism of dispersal, support the idea that coastal rivers act as a linear array of islands of divergence, and that southern lineages will be the most diverged (April et al., 2013; Carvajal-Quintero et al., 2019; Dias et al., 2014; Eckert et al., 2008; Le Corre & Kremer, 1998; Sepkoski & Rex, 1974). During the Pliocene, mean sea levels dropped from 50 to 80 m above current sea level to an estimated 80 to 100 m below (Haq, Hardenbol, & Vail, 1987; Riggs, 1984), which would have facilitated lateral transfer from the presumed origin of the genus in the Mississippi-Missouri valley moving southwest (e.g. Nagle & Simons, 2012). After lateral transfer during the late Pliocene and early Pleistocene, sea levels began to rise 10 to 20 m above present day levels, thus severing the coastal connections (Haq et al., 1987; Riggs, 1984). Subsequent glacial advances (potentially up to 16; Martinson et al., 1987) and periods of lowered sea levels during the Pleistocene, would have allowed for ichthyofaunal transfer

between lower reaches of basins which were geographically closer (e.g. Brazos and Colorado rivers, but likely not the Brazos River and Rio Grande; Conner & Suttkus, 1986). This is supported by the genetic relationships in this study in that the lineages in the most distant and southwestern drainages (e.g. Rio Grande and Guadalupe-San Antonio basin) are the most genetically diverged relative to lineages in basins which are geographically closer (e.g. Colorado, Brazos, Red River) to the presumed epicenter of the genus (e.g. Mississippi River). Thus, it is likely that the ancestral lineage was able to disperse laterally along the coastal drainages, reached the Rio Grande and Guadalupe-San Antonio rivers, then the connections were severed but the lineages were able to persist and diverge in isolation.

The lineages found in the Colorado River are the only ones found in multiple drainages, suggesting the Colorado River may have a more complex geographical history than other Gulf Slope rivers or its central location relative to the drainages in this study provided more opportunities for faunal transfer. *Macrhybopsis marconis* is found in the Colorado and Guadalupe-San Antonio rivers, and a lineage of *M. hyostoma* is found in the Brazos and Colorado rivers. The genetic relationships of these lineages suggest that they diverged in one river, potentially the Guadalupe River for *M. marconis*, which was much larger during the Pliocene (Galloway et al., 2011), and the Brazos for *M. hyostoma*, then were able to disperse into the other river, likely via river capture or during the late Pleistocene when the area was much more wet (Conner & Suttkus, 1986; Wilkinson & Basse, 1978). *Macrhybopsis hyostoma* could have dispersed between the Brazos and Colorado rivers as recently as 9,000 to 13,000 years ago, as those waterways were connected with an ancient river known as the Galveston River (Conner & Suttkus, 1986).

Additionally, the Colorado River has similar ichthyofauna composition to the Brazos River (87% community similarity) and San Antonio basin (83%), suggesting connections allowing for relatively recent ichthyofaunal transfer, whereas it is less similar to the Rio Grande (61%; Conner & Suttkus, 1986; Craig et al. unpublished data). Thus, it is likely *M. marconis* and BrCol *M. hyostoma* diverged within a single drainage and were able to transfer via river capture or coastal connections (Albert, Craig, Tagliacollo, & Petry, 2018; Burridge, Craw, & Waters, 2006; Conner & Suttkus, 1986; Nagle & Simons, 2012; Osborne et al., 2016).

The rest of the lineages in this study show lower levels of genetic structure and are found within the Mississippi River basin. These lineages are relatively young and likely diverged via dispersal followed by isolation, which could have occurred throughout the Pleistocene. However, the upstream lineages, *M. australis*, *M. tetranema,* Clear Fork *M. hyostoma,* and even Red *M. hyostoma,* present an interesting case worth investigating further. These four lineages are found in systems which contain a more downstream *M. hyostoma* lineage, which are generally more genetically similar to the endemic lineages than other *M. hyostoma* (Echelle et al., 2018; Gilbert et al., 2017; Underwood et al., 2003). This suggests that the upstream lineages and the *M. hyostoma* lineage within each basin share a most recent ancestor, or the upstream endemic diverged from the downstream *M. hyostoma* lineage. As it stands, *M. australis* and *M. tetranema* have clear divergence from the other lineages, as they show unique morphology (Eisenhour, 2004) and show divergence in all analyses herein and in other work (Echelle et al., 2018; Underwood et al., 2003). However, Clear Fork *M. hyostoma* and Red *M. hyostoma* show mixed ancestry and incomplete divergence in some analyses, although they split out in

the phylogenetic analysis. Thus, these two lineages are younger and are currently diverging from the Brazos Colorado *M. hyostoma* and the Mississippi *M. hyostoma*, respectively, and are limited to the upper reaches of the drainage, similar to *M. australis* and *M. tetranema*. This is somewhat unsurprising as upstream reaches of rivers are less open to migrants or new species, even when considering historical connections, and are therefore more isolated than downstream reaches (Burridge, Craw, Jack, King, & Waters, 2008; Carvajal-Quintero et al., 2015; Schmera et al., 2018). As these endemic lineages are generally confined to the upstream reaches, the divergence could have started due to drying during a glacial retreat and subsequently been reinforced due to unidirectional gene flow in rivers or local adaptation to more extreme conditions (Carvajal-Quintero et al., 2015; Datry et al., 2016; Sotola et al., 2019). However, *M. australis* does hybridize with Red *M. hyostoma*, indicating these lineages are not completely reproductively isolated (Sotola et al., 2019), which may also be the case for other lineages; further research is warranted.

Overall, this study shows that river connectivity played a major role in the evolutionary history of the *Macrhybopsis* complex, and that the diversity of these riverine minnows is likely larger than currently recognized. The *Macrhybopsis* genus is widespread throughout the United States, with the majority of the genus occurring in the Mississippi River basin. Present-day distributions across multiple river basins were achieved via river connections which no longer exist. Thus, the present work supports that historical connectivity is the main determinant of range size (Carvajal-Quintero et al., 2019). If the ancestral Mississippi River basin is the epicenter and origin of *Macrhybopsis* complex (Eisenhour, 2004), all drainages would have at one point had a

connection allowing for dispersal, followed by periods of isolation allowing for divergence. This general pattern suggests dispersal during times when the Mississippi River was much larger, sea level was much lower, and during wet periods of the Pleistocene (Conner & Suttkus, 1986; Galloway et al., 2011; Haq et al., 1987; Riggs, 1984) followed by periods of isolation allowing for vicariance. This has been shown to be a powerful mechanism in the evolutionary history of fishes, with dispersal followed by vicariance responsible for numerous fish species (Burridge et al., 2006; Leroy et al., 2019; Machado, Galetti, & Carnaval, 2018; Mayden, 1988; Pérez-Rodríguez, Domínguez-Domínguez, Mar-Silva, Doadrio, & Pérez-Ponce de León, 2016). Additionally, this work supports results showing that the more southern lineages are more diverged from northern lineages, indicating they have been persisting longer in isolation without being extirpated by glaciation or reconnected to other river basins allowing gene flow (April et al., 2013). Lastly, the lineages in this study conform to the stepping-stone model of dispersal and the central-marginal hypothesis where they more distant and isolated drainages will contain the more diverged lineages (Eckert et al., 2008; Le Corre & Kremer, 1998; Sepkoski & Rex, 1974).

Further support for the roles of dispersal and vicariance in producing current patterns of distribution and divergence is the fact that within the *Macrhybopsis* genus, there appear to be high amounts of endemism to single drainages, both in this and previous studies (Echelle et al., 2018; Gilbert et al., 2017; Underwood et al., 2003). This study reported genetic divergence within the *M. hyostoma* species which are similar to or higher than between other recognized species (e.g. *M. australis* and *M. tetranema*). Other studies on *Macrhybopsis* have reported similar findings where populations in separate

drainages appear to be relatively highly diverged (Echelle et al., 2018; Gilbert et al., 2017; Underwood et al., 2003). The high amounts of endemism in *Macrhybopsis* and fishes in Gulf Slope drainages suggest that these drainages acted as glacial refugia during the dry periods of the glacial cycles. During the glacial cycles of the Pleistocene, large parts of the country were dry during a glacial retreat, thus the river systems which stayed wet were able to support fishes and other aquatic fauna (Jetz, Rahbek, & Colwell, 2004; Oberdorff, Lek, & Guegan, 1999; Tedesco et al., 2012). These glacial cycles could last for thousands of years which would have allowed for divergence between basins, and have been suggested as an explanation for the large amount of endemism in Gulf Slope drainages (Craig, Kollaus, Behen, & Bonner, 2016).

Recent work on a wealth of taxonomically diverse organisms has revealed large amounts of undiscovered biodiversity, including within fishes (Alter, Munshi-South, & Stiassny, 2017; Bartáková et al., 2019; Dufresnes et al., 2020, 2019; Fletcher et al., 2019; Healey et al., 2018; McGaughran, Terauds, Convey, & Fraser, 2019; Melo, Ochoa, Vari, & Oliveira, 2016; Walker et al., 2009). Further, half of the publications which used the term “cryptic species” have been published since 2014 (Dufresnes et al., 2020) suggesting either (or both) an increased effort in evolutionary lineage designation or more efficient and effective techniques. This work and other recent work on *Macrhybopsis* and other riverine fishes support this idea that there is a large amount of undiscovered biodiversity. Uncovering cryptic diversity and understanding the patterns and processes that led to this amount of biodiversity in riverine fishes and defining evolutionary lineages is a vital process in evolutionary biology, conservation, and management efforts. This work has demonstrated that the biodiversity of small bodied riverine fishes is likely

much higher than currently recognized. Additionally, this work demonstrates that historical connections allowing for dispersal, followed by periods of isolation allowing for vicariance have greatly influence the amount of genetic structure seen in riverine fishes.

Table 2.1. Recognized species, rivers where captured, identified lineages according to this study (see results), abbreviation, total sample size (N total), final sample size per location (N final), and downsampling sample size (N per lineage).

Species	River(s)	Lineage	Abbreviation	N total	N final	N per lineage
M. aestivalis	Rio	*M. aestivalis*		30	25	25
M. australis	Pease	*M. australis*		97	10	25
	Red			83	5	
	Wichita			114	10	
M. hyostoma (1)	Lower Brazos	Brazos Colorado *M. hyostoma*	BrCol *M. hyostoma*	59	12	24
	Colorado			29	12	
M. hyostoma (2)	Clear Fork Brazos	Clear Fork *M. hyostoma*	C.F. *M. hyostoma*	12	12	24
	Upper Brazos			40	12	
M. hyostoma (3)	Mississippi	Mississippi *M. hyostoma*	M.R. *M. hyostoma*	20	19	19
M. hyostoma (4)	Red	Red *M. hyostoma*	R.R. *M. hyostoma*	177	24	24
M. marconis	Colorado	*M. marconis*		10	10	24
	Guadalupe			55	7	
	San Antonio			12	7	
M. tetranema	Canadian	*M. tetranema*		38	24	24

Table 2.2. Procrustes correlation coefficients (mean and standard deviation [SD]) of PCA plots between subsampled SNPs (N = 50, 100, 250, 500, and 1000) and the full dataset (N = 20,376 SNPs).

N SNPs	Procrustes Correlation	Standard Deviation
50	0.895	0.005
100	0.927	0.004
250	0.972	0.002
500	0.987	0.001
1000	0.993	0.0002

Table 2.3. Scenarios tested in DIYABC (N = 16), scenario name, sources, and figures in the present study, if applicable.

Scenario	Name	Source	Figure
1	Simultaneous divergence	Null model of simultaneous divergence	
2	Phylogenetic relationships	SNAPP analysis	Figure 2.7
3	Entropy relationships	Entropy analysis	Figure 2.5
4	Vicariance	Eisenhour 2004	
5	Vicariance	Eisenhour 2004	
6	Vicariance	Eisenhour 2004	
7	Vicariance	Eisenhour 2004	
8	Vicariance	Eisenhour 2004	
9	Genetic differentiation	G_{ST}	Figure 2.2
10	Underwood phylogenetic	Underwood et al. 2003	
11	Stepping-stone/Central-marginal hypothesis	Sepkoski & Rex, 1974; Eckert et al., 2008	
12	Stepping-stone/Central-marginal hypothesis	Sepkoski & Rex, 1974; Eckert et al., 2008	
13	Echelle Figure 2	Echelle et al. 2018	
14	Echelle Figure 3	Echelle et al. 2018	
15	Stepping-stone/Central-marginal hypothesis	Sepkoski & Rex, 1974; Eckert et al., 2008	
16	Stepping-stone/Central-marginal hypothesis	Sepkoski & Rex, 1974; Eckert et al., 2008	

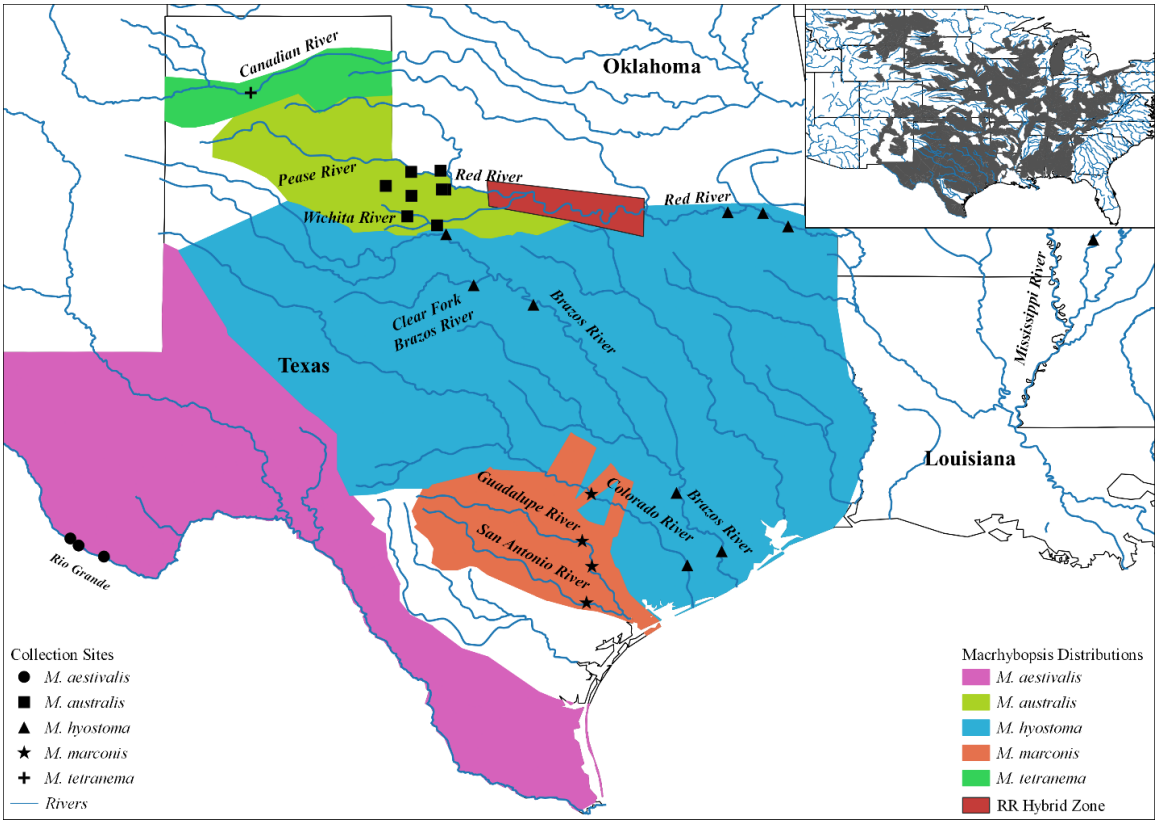


Figure 2.1. Map of collection locales for each lineage. Plus signs (+) are *M. tetranema* from Canadian River, squares are *M. australis* from the Red, Pease, and Wichita rivers, triangles are *M. hyostoma* from the Red, Mississippi, Brazos, Clear Fork, and Colorado rivers, stars are *M. marconis* from the Colorado, Guadalupe, and San Antonio rivers, and circles are *M. aestivalis* from the Rio Grande. Shaded colors are approximate distributions of recognized species within Texas: green is the Canadian River *M. tetranema*, yellow is Red River basin *M. australis*, red is the Red River hybrid zone of *M. australis* and *M. hyostoma*, blue is Red, Brazos, Clear Fork, and Colorado rivers *M. hyostoma*, orange is Colorado, Guadalupe, and San Antonio rivers *M. marconis*, and purple is Rio Grande *M. aestivalis*. Inset map at the top right is the distribution of the *Macrhybopsis* genus within the contiguous U.S.A.

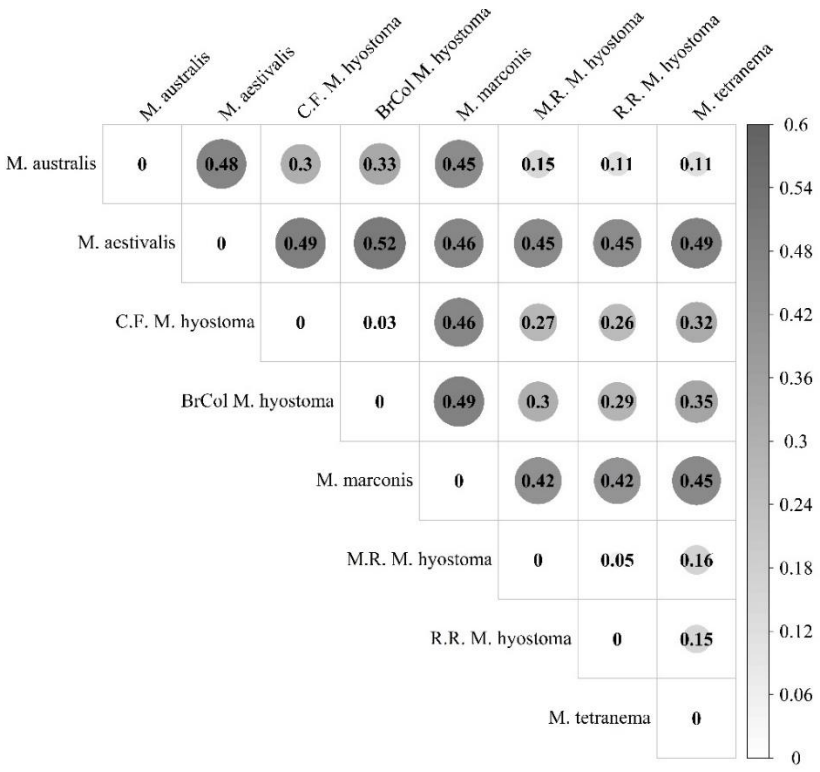


Figure 2.2. Heat map of G_{ST} calculated from allele frequencies, including the actual G_{ST} values. Abbreviations for each lineage (if necessary) are in Table 2.1.

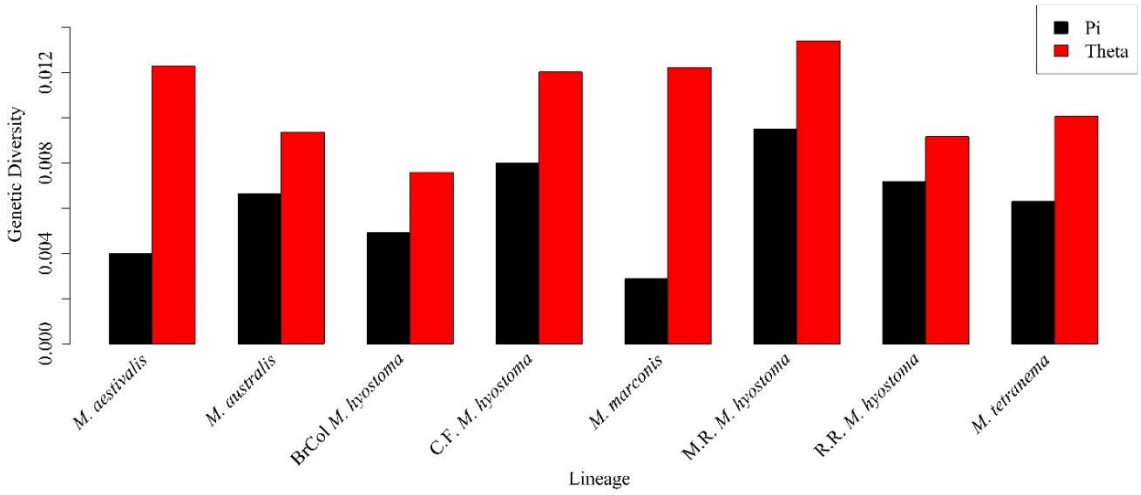


Figure 2.3. Genetic diversities, pi (black bar) and theta (red bar), as measured for each lineage. Abbreviations for each lineage (if necessary) are in Table 1.Heat map of G_{ST} calculated from allele frequencies, including the actual G_{ST} values. Abbreviations for each lineage (if necessary) are in Table 2.1.

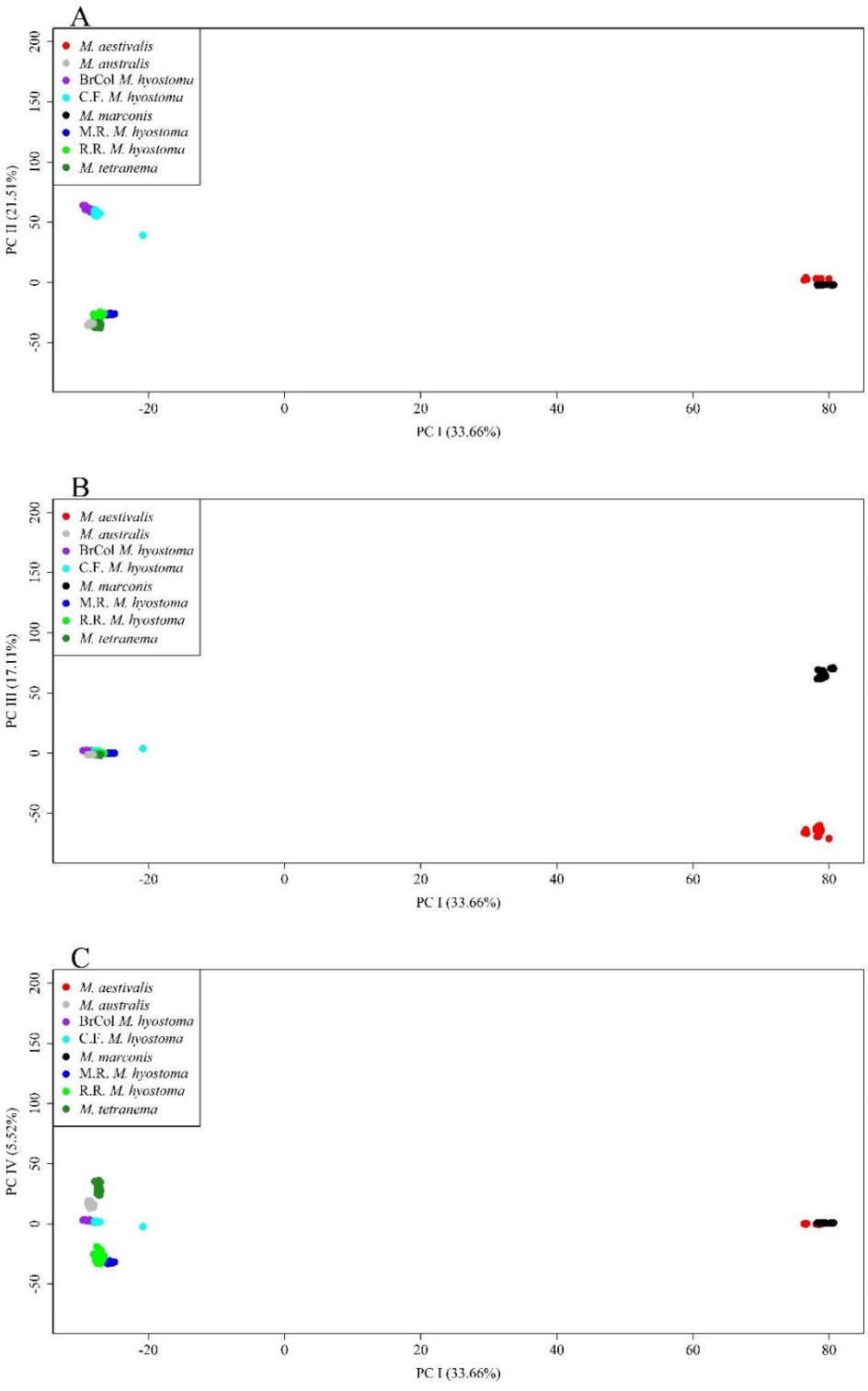


Figure 2.4. PCA of genetic differentiation of all individuals from each lineage. PCI and PCII are plotted in A, PCI and PCIII are plotted in B, and PCI and PCIV is plotted in C. Colors represent different lineages. See Table 2.1 for abbreviations.

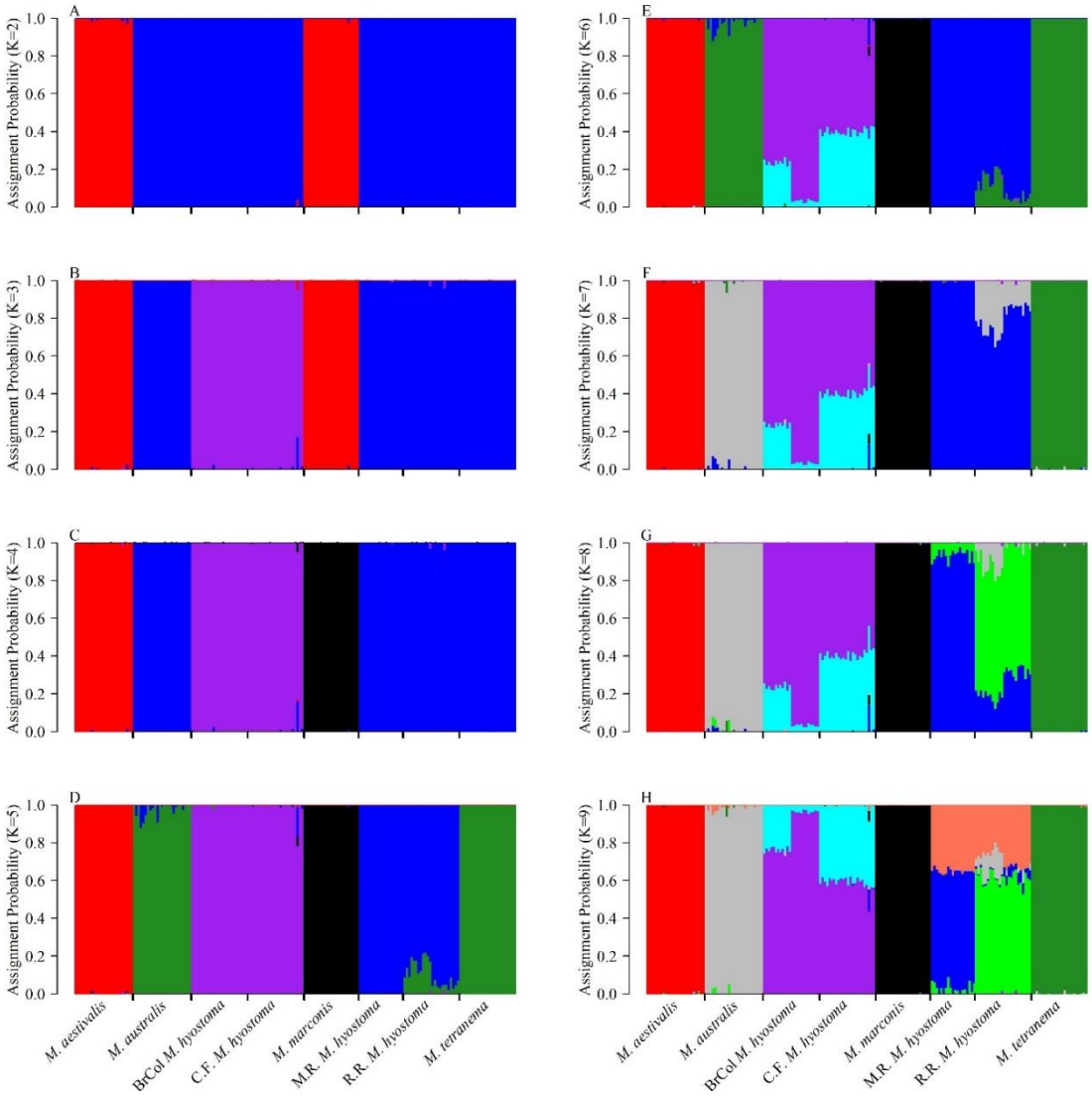


Figure 2.5. Entropy plots for k = 2 (A) through k = 9 (H) for all lineages throughout Texas. See Table 2.1 for abbreviations.

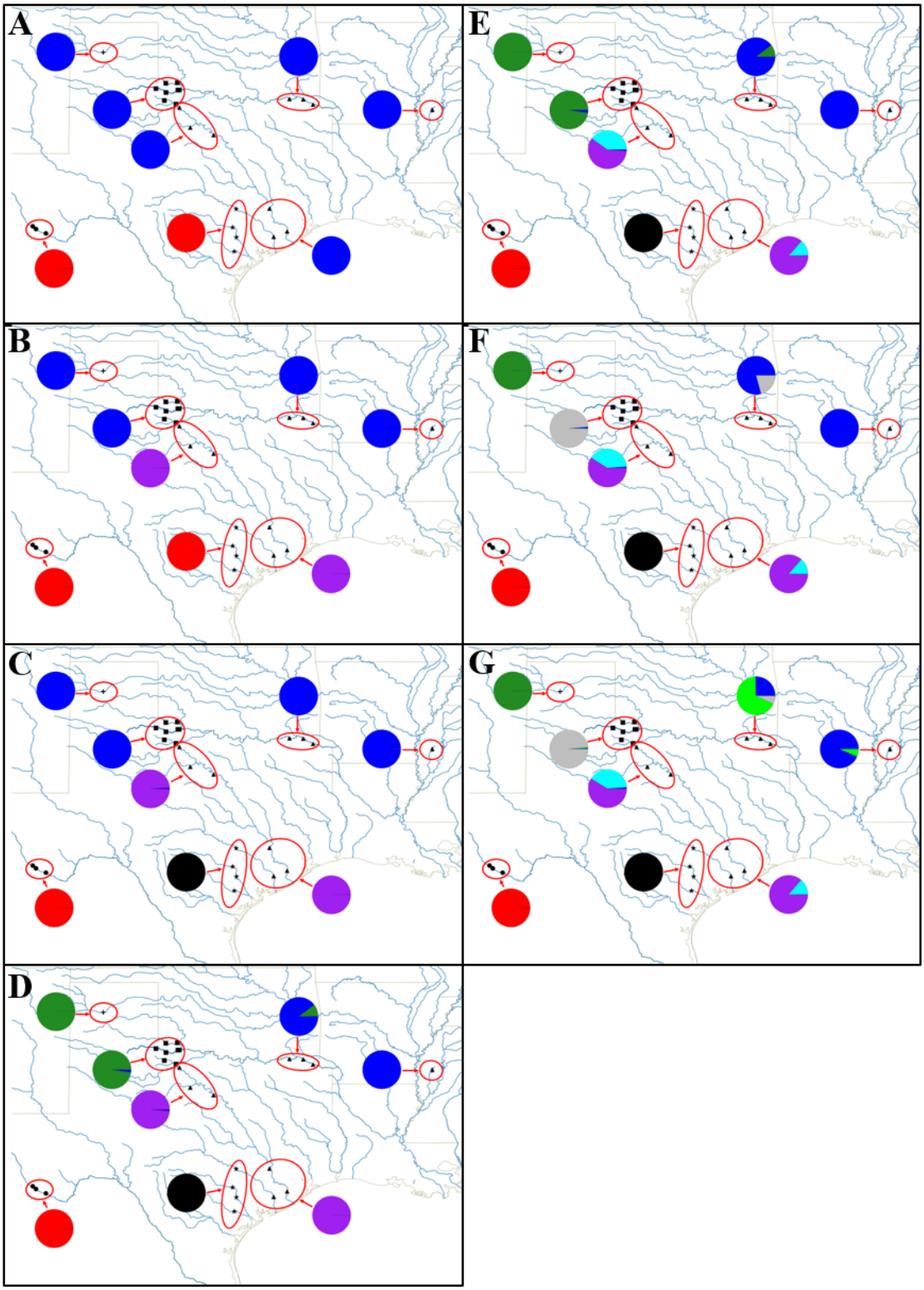


Figure 2.6. Entropy plots as pie graphs showing average assignment for members of each lineage for k = 2 (A) through k = 8 (H) for all lineages throughout Texas.

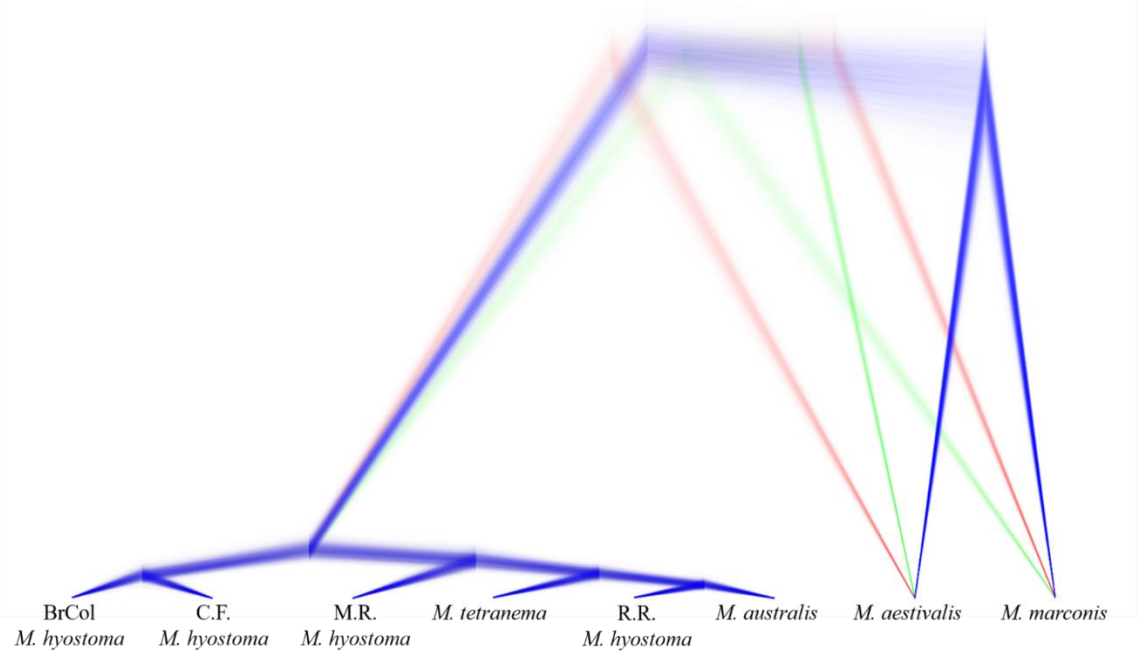


Figure 2.7. Phylogenetic relationships of each lineage in this study made from SNAPP analysis via Beast. See Table 2.1 for abbreviations.

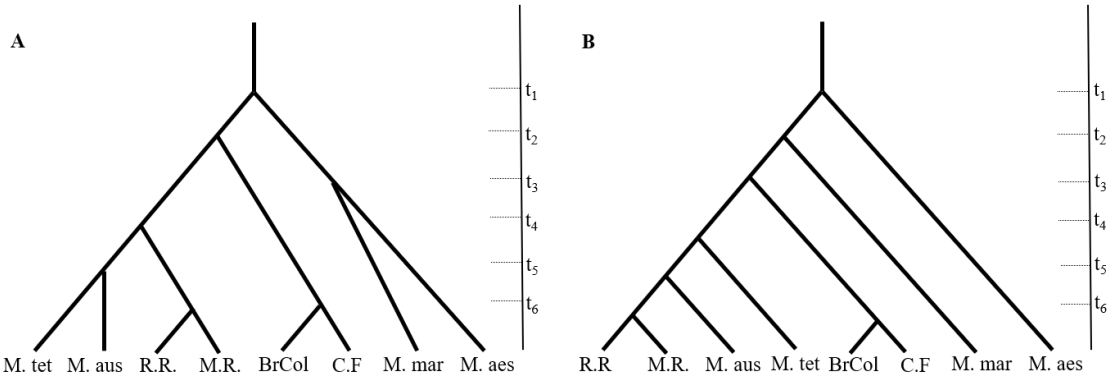


Figure 2.8. Approximate Bayesian Computation models which are most probable as estimated from DIYABC. These are the most likely scenarios calculated via the direct (A) and logistic (B) methods. See Appendix 2.1 for all tested scenarios.

III. POPULATION-LEVEL RESPONSES OF FRESHWATER MUSSELS TO FLOODS IN A SOUTHWESTERN U.S.A. RIVER USING MARK-RECAPTURE SAMPLING

Abstract

Floods can directly affect riverine organisms by displacing them, and population-level responses to floods can vary depending on flood magnitude and organism mobility. Benthic organisms can resist displacement until substrates become unstable, whereas mobile organisms are more resistant. Freshwater mussels are benthic organisms with low mobility and limited research on their population-level responses to floods. This study provides novel insight to population-level responses of mussels to large floods (>500 m^3/s). Population dynamics (i.e., detection probability, abundance, and apparent survival) in a closed robust design framework were estimated for four freshwater mussel species (*Cyclonaias petrina*, *C. pustulosa*, *Amblema plicata*, and *T. verrucosa*) from 2017 to 2019 at two sites (upper and lower sites) within riffle habitats in the Colorado River, Texas, USA. During sampling, a flood occurred at both sites, each of which was in the 99^{th} percentile of historical flows at their respective gages yet were of different magnitudes. Each species was affixed with shellfish tags, with *C. petrina* and *C. pustulosa* also being affixed with PIT tags. This allowed direct comparison of population-level estimates from data collected using each tag type. There were site- and species-specific differences in estimated abundances and apparent survival during periods with the floods. Estimated abundances of PIT-tagged *C. petrina* and *C. pustulosa* were reduced 31 to 45% with the lesser flood magnitude (1,283 m^3/s). Estimated abundances of PIT-tagged *C. petrina* and *C. pustulosa* were reduced 73 to 80% with the greater flood

magnitude (4,332 m^3/s). Reductions in estimated abundances were not detected for shellfish-tagged *A. plicata* and *T. verrucosa*. There were no differences in apparent survival at the upper site, while initially high apparent survival at the lower site was reduced during the interval with the flood. Population estimates from shellfish tags were generally lower and more variable than estimates from PIT tags indicating underestimation of population parameters. Floods reduced the abundance of two species within riffle habitats at the two sites. Large floods, therefore, affect populations dynamics of mussels, but the fate of the displace mussels are unknown and reach-scale effects are unknown at this time. This study adds to the growing body of knowledge on vagile aquatic organism's resistance to large floods, although quantification of resiliency is needed to fully understand long-term fitness of mussels with large floods.

Introduction

Resistance and resiliency of riverine organisms to flooding are often-studied topics in population and community ecology (Flecker & Feifarek, 1994; Franssen et al., 2006; Grimm & Fisher, 1989; Maltchik & Pedro, 2001; Power & Stewart, 1987; Robinson, 2012). Floods, generally defined as occurrences of water in usually dry areas (Jonkman & Kelman, 2005), indirectly affect riverine communities by altering physical (Peters, Caissie, Monk, Rood, & St-Hilaire, 2016) and chemical (Talbot et al., 2018) components of lotic systems. They also directly affect communities by displacing organisms (Cobb, Galloway, & Flannagan, 1992), and may have a variety of effects on ecosystem functions and services (Talbot *et al.*, 2018 and references therein). Additionally, floods are considered essential components of the flow regime, maintaining ecological integrity of riverine communities (Poff et al., 1997). Population-level

responses of riverine organisms to floods vary widely, depending on several factors including flood magnitude, organismal biology (e.g. mobility), and instream habitats. Generally, populations are more resistant to small floods than large floods, though the effects will depend on stream geomorphic features and hydraulic forces (Robinson, 2012). At the individual level, mobile organisms (e.g., fish and mammals) are more resilient to displacement effects (Crandall, Hayes, & Ackland, 2003) due to their ability to escape or find appropriate refuge, whereas less mobile, sessile organisms (e.g., rooted plants and some invertebrates), are particularly sensitive to flooding and resist displacement unless substrates become unstable (Cobb et al., 1992).

Information about responses to floods is reported for a few populations of freshwater mussels, a group of benthic organisms with limited mobility. A flood in one southwest (USA) desert river (maximum daily flow: 26 m^3/s, percentile: 99^{th}, median flow: 0.20 m^3/s, drainage area: 890 km^2; USGS Station 08405500) had no detectable effect on mussel population dynamics (e.g., survival, growth) over a 15-year period (Inoue, Levine, Lang, & Berg, 2014). Moreover, the flood event was considered beneficial for mussel survival because it displaced fine sediments that accumulated during low flow periods (Inoue et al., 2014). Floods in two northeastern (USA) creeks, Tonawanda Creek (maximum daily flows: 163 m^3/s, percentile: 99^{th}, median flow: 6.8 m^3/s, drainage area: 900 km^2; USGS Station 04218000) and French Creek (maximum daily flows: 479 m^3/s, percentile: 99^{th}, median flow: 38 m^3/s, drainage area: 2,040 km^2; USGS Station 03023100), had no detectable effects on mussel survival, occurrences, or composition over two decades, despite bed sediment mobilization occurring during more frequent floods (<2 year intervals) (Sansom et al., 2018). Conversely, a flood described

as a 100-year flood event on an ungaged, upland river in Scotland (Hastie, Boon, Young, & Way, 2001) was a potential conservation concern for a mussel population, where an estimated 50,000 mussels, representing 4 to 8% of the total population, were displaced, stranded, and died. In other studies, the ability of freshwater mussels to resist floods are reported to depend on substrate stability and hydraulic variables related to substrate stability such as shear stress (Allen & Vaughn, 2010; Gangloff & Feminella, 2007; Morales, Weber, Mynett, & Newton, 2006; Randklev, Hart, Khan, Tsakiris, & Robertson, 2019; Strayer, 1999; Zigler, Newton, Steuer, Bartsch, & Sauer, 2008), habitat type (Meador, Peterson, & Wisniewski, 2011), channel geomorphology (Gangloff & Feminella, 2007), and shell morphology, behavior, and life-history strategies of the mussels (Allen & Vaughn, 2009; Goodding, Williams, Ford, Williams, & Ford, 2019; Randklev et al., 2019). To date, empirical studies that directly assess effects of floods on mussel population dynamics in large rivers are lacking, particularly in relation to large floods (e.g. >500 m^3/s).

In spring and summer 2017, two mark-recapture sites located in the upper and lower Colorado River, Texas (USA) —henceforth referred to as "upper site" and "lower site"— were established within riffle habitats to quantify population dynamics of mussel populations. Target species included *Cyclonaias petrina*, an endemic state-listed threatened species and a candidate species for listing by U.S. Fish and Wildlife Service, and *C. houstonensis*, another candidate species for listing, which was later synonymized with *C. pustulosa* (Johnson et al., 2018). Additional target species were two common mussel species, *Tritogonia verrucosa* and *Amblema plicata*. Mark-recapture studies commonly use shellfish tags (Inoue et al., 2014; Newton, Zigler, & Gray, 2015;

Wisniewski, Shea, Abbott, & Stringfellow, 2013) to estimate population dynamics such as abundances, immigration, emigration, and survival. However, burrowing tendencies of mussels can make them difficult to recapture in tactile surveys, leading to underestimates in population parameters (Strayer & Smith, 2003; Wisniewski et al., 2013). Thus, to improve estimates of population parameters in this study and to additionally assess the effectiveness of shellfish tags, passive integrated transponder tags (PIT tags) were used on candidate species, in addition to shellfish tags on all target species in a closed robust design framework. During the study period, in August 2017, precipitation from Hurricane Harvey inundated the lower site with a peak flow of 4,332 m^3/s (percentile: 99^{th}, median flow: 44 m^3/s, drainage area: 110,000 km^2; USGS Station 08161000). In October 2018, precipitation from a frontal boundary inundated the upper site with a peak flow of 1,283 m^3/s (percentile: 99^{th}, median flow: 5.6 m^3/s, drainage area: 51,000 km^2; USGS Station 08147000). Both floods were classified as greater than one per five-year events (Buzan *et al.*, 2011).

The purpose of this study was to opportunistically assess population-level responses of four mussel species within riffle habitats following large (99^{th} percentile, greater than one per five-year events) floods. Objectives were to quantify population-level responses (detection probability, estimated abundance, and apparent survival) of four mussel species (*C. pustulosa, C. petrina, T. verrucosa,* and *A. plicata*) using shellfish tags, in addition to PIT tags on two of the species (*C. petrina* and *C. pustulosa*). Estimation of population-level responses using shellfish tags and PIT tags on *C. petrina* and *C. pustulosa* enabled comparison of estimated population-level responses between two tag types. The latter objective will improve methods for estimating mussel responses

as future work continues in assessing effects of flood events. Floods were expected to reduce abundance and apparent survival of the four mussel species with greater reductions at a peak flow of 4,332 m^3/s at the lower site than a peak flow of 1,283 m^3/s at the upper site unless substrate differences between sites mediated the decreases in abundance and apparent survival.

Methods

Two sites on the Colorado River with high densities of mussel species (Ruppel 2019) were chosen as mark-recapture locations. The upper site, located in the Colorado River near San Saba, Texas, was a riffle with mixture of cobble (60%), sand (25%), and gravel (15%) on the standard Wentworth scale (Wentworth, 1922). Water quality parameters were measured with a multiprobe meter (YSI-85) during sampling. Water temperature ranged from 15.1 to 29.6°C, dissolved oxygen ranged from 7.2 to 10.5 mg/l, and specific conductance ranged from 501 to 711 µS/cm during the study. The lower site, located in the lower Colorado River near Columbus, Texas, was a riffle with predominately cemented sandstone (70%) with interstitial pockets of sand (20%) and gravel (10%). Water temperature ranged from 20.8 to 31.7°C, dissolved oxygen ranged from 7.5 to 10.9 mg/l, and specific conductance ranged from 574 to 712 µS/cm during the study.

Robust design mark-recapture methods (Pollock, 1982; Nichols & Pollock, 1990) were used to estimate detection probability, abundance, and apparent survival of freshwater mussels. Robust design methods consist of primary and secondary periods, where populations are assumed to be closed (i.e. no mortality or migration), and intervals, defined as the time between primary periods, where the populations are assumed to be

open (i.e. mortality or migration can occur). At the upper site, mussels were initially captured and tagged in June 2017 and subsequently sampled during five primary periods over a three-year span (August and November 2017, April and August 2018, and April 2019). At the lower site, mussels were initially captured and tagged in March 2017 and sampled during five subsequent primary periods over two years (April, August, and November 2017; April and August 2018). Primary periods, consisting of three secondary periods (i.e. sampling events), were separated by three to four months; however, there were instances in which sampling had to be delayed several weeks for high flows to subside. Secondary periods were separated by about 24 hours. In total, there were five primary periods and four intervals at each sampling site.

For initial tagging and during subsequent primary and secondary periods, a 300-m^2 rectangular area was delineated within a riffle habitat at each site. The four corners were georeferenced so that the same area could be delineated during subsequent visits. During initial sampling, survey crews spread evenly across the downstream boundary and searched for mussels visually and tactilely moving upstream while crawling, floating, or snorkeling. Detected mussels were removed and placed into mesh bags kept in the river. Upon completion of the survey, mussels were taken to a central processing station on the riverbank and identified morphologically to species. Mussels were then affixed with two laminated vinyl shellfish tags (Floy®) on each valve. For *C. petrina* and *C. pustulosa*, a PIT tag (Biomark ®) was also affixed to a valve. Cyanoacrylic glue (Loctite Gel Control Super Glue®) was used to affix tags to the mussel valves (Young & Isely, 2008; Ashton, Tiemann & Hua, 2017). Mussels were returned to the 300-m^2 rectangular area and placed in substrates with their posterior end in an upright position. For subsequent primary and

secondary period sampling, the 300-m^2 rectangular areas were surveyed using a Biomark reader to locate PIT tagged individuals. After scanning, mussels were visually and tactilely captured, tagged, and returned as during initial tagging. For previously tagged mussels, the unique tag number per recaptured individual was recorded. Average person hours (p-h; calculated as total search time divided by number of people) ranged from 25.3 to 73.5 p-h at the upper site, and 4.6 to 36.0 p-h at the lower site. Mussel species with more than 15 total individuals captured were included in analyses. This included *C. petrina*, *C. pustulosa*, and *T. verrucosa* at the upper site, and *C. petrina*, *C. pustulosa*, and *A. plicata* at the lower site.

Discharge at the upper site, measured from USGS gage 08147000, ranged from 0.06 to 1,283 m^3/s throughout the duration of the study (Figure 3.1A). Discharge at the lower site, measured from USGS gage 08161000, ranged from 13 to 4,332 m^3/s (Figure 3.1B; Table 3.1). The flood at the upper site occurred during interval four, and the flood at the lower site occurred during interval two (Table 3.1). Median daily flow at the upper site (period of record: 1915–2017) was 5.6 m^3/s with a maximum peak flow of 5,409 m^3/s in 1938. Median daily flow at the lower site (period of record: 1915–2017) was 44 m^3/s with a maximum peak flow of 4,642 m^3/s in 1935.

A Bayesian closed robust design model (Pollock, 1982; Nichols & Pollock, 1990) was used to estimate apparent survival and abundance while accounting for imperfect detection separately for each species and site using methods and code modified from Riecke *et al.* (2018). Specifically, detection probability (p_t) was estimated as the probability an individual available for detection during time *t* was in fact detected during at least one secondary occasion, abundance (N_t) was estimated as the number of

individuals in the study area during each primary period, and apparent survival (ϕ_t) was estimated as the probability an individual in the population at time *t* survived to time *t + 1* and did not permanently emigrate from the study area. Uninformative priors were used for all parameters; for detection probability and apparent survival a uniform distribution on the interval [0,1] was used. The probability (Pr) of losing one PIT tag or shellfish tag (t_1) or both tags (t_2) was estimated as:

$$\Pr[t_1] = \frac{l}{N}(1 - \frac{l}{N})$$

$$\Pr[t_2] = t_1^2$$

where *l* is the number of mussels observed with tag loss and *N* is the total number of tagged mussels (Reinert *et al.*, 1998; Meador *et al.*, 2011). In this study, the probability of losing a PIT tag was 0.053, the probability of losing a shellfish tag was 0.024, the probability of losing two shellfish tags was 0.00058, and the probability of losing a PIT and shellfish tag was 0.00127. The probability of not losing tags (1 - t_i) was calculated and was used as a constant multiplier on apparent survival and estimated abundances in each model for each species depending on which type of tags were on each individual to account for tag loss in these estimates.

All model parameters were estimated using Markov Chain Monte Carlo methods in JAGS (Plummer 2003) written in the BUGS language, through R (R Core Team 2019) with the R2jags package (Su & Yajima, 2015, see Appendix 3.6 for model code). A total of 15,000 iterations were used with a burn-in period of 5,000 iterations, and a thinning rate of 10 for each of three Markov chains to ensure sufficiently large effective sample

sizes (range 140 to 3,000; mean of 2,385 for all parameters). Convergence was confirmed using visual inspection of trace plots and ensuring that the Gelman-Rubin statistic (Gelman & Rubin, 1992) was less than 1.10 for all parameters. All estimates are presented as posterior medians with 95% credible intervals (CRI) presented in Appendix 3.1 through 3.4.

Parameter estimates were compared between each primary period (p_t and N_t) or interval (ϕ_t) and the one directly preceding it in time to detect differences over time (i.e., t_1 and t_2 were compared, but not t_1 and t_3). For each MCMC iteration, the posterior estimate from time t was subtracted from the posterior estimate for time $t + 1$ to represent increases (+) or decreases (-) from one time to the next. Credible intervals (95% CRIs) were estimated around the differences, and if the CRI excluded zero, then the difference in estimates between periods was considered statistically significant.

Posterior estimates of detection probability, abundances, and apparent survival of *C. petrina* and *C. pustulosa* between PIT tag and shellfish tag data were compared using methods described above. The posterior estimates of the shellfish tag data were subtracted from the posterior estimates of the PIT tag data, and a 95% CRI for the difference was estimated. If the credible intervals excluded zero, the difference was interpreted as statistically significant. Negative values indicated that the parameters from shellfish tag data were underestimated relative to those from PIT tag data, while positive values indicated that they were overestimated.

Results

At the upper site, PIT tags were affixed to 442 *C. petrina* and 15 *C. pustulosa* among five primary periods with the 1,283 m^3/s flood occurring between primary periods

4 and 5 and during interval 4 (Table 3.2). For *C. petrina*, the probability of detecting a PIT-tagged individual ranged from 0.693 to 0.972 (Figure 3.2; Appendix 3.1 & 3.2). Estimated abundances increased from 153 to 363 pre-flood (primary periods 1 - 4) and decreased to 249, a 31% decrease, after the flood (primary period 5; Table 3.3). Differences were not detected in apparent survival between intervals. For *C. pustulosa*, the probability of detecting a PIT-tagged individual ranged from 0.789 to 0.973. Estimated abundances generally increased from 6 to 11 pre-flood (primary periods 1 - 4) and decreased to 6, a 45% decrease, after the flood (primary period 5). Differences were not detected in apparent survival between intervals.

At the upper site, shellfish tags were affixed to 442 *C. petrina*, 15 *C. pustulosa*, and 217 *T. verrucosa* across the five primary periods. For *C. petrina*, the probability of detecting a shellfish-tagged individual ranged from 0.335 to 0.958 (Figure 3.3; Appendix 3.3 & 3.4). Estimated abundances ranged from 149 to 240 pre-flood (primary periods 1 - 3); differences in estimated abundances were not detected after the flood (primary periods 4 and 5). Apparent survival decreased from 0.978 during interval 2 to 0.706 during interval 3 pre-flood. For *C. pustulosa*, the probability of detecting a shellfish-tagged individual ranged from 0.639 to 0.958 among primary periods. Estimated abundances ranged from 4 to 7 pre-flood (primary periods 1 - 4); differences in estimated abundances were not detected after the flood (primary periods 4 and 5). For *T. verrucosa*, the probability of detecting a shellfish-tagged individual ranged from 0.339 to 0.970 among primary periods. Estimated abundances ranged from 62 to 110 pre-flood (primary periods 1 - 3); differences in estimated abundances were not detected after the flood (primary periods 4 and 5). Differences were not detected in apparent survival between intervals for

C. pustulosa or *T. verrucosa*.

At the lower site, PIT tags were affixed to 120 *C. petrina* and 308 *C. pustulosa* among five primary periods with the 4,332 m^3/s flood occurring between primary periods 2 and 3 and during interval 2 (Table 3.2). For *C. petrina*, the probability of detecting a PIT-tagged individual ranged from 0.160 to 0.999 (Figure 3.4; Appendix 3.1 & 3.2). Estimated abundances increased from 69 to 101 pre-flood (primary periods 1 and 2) and decreased to 28, a 73% decrease, after the flood (primary period 3). Apparent survival was 0.933 pre-flood (interval 1), decreased to 0.237 during the flood (interval 2), and increased to 0.796 after the flood (interval 3). For *C. pustulosa*, the probability of detecting a PIT-tagged individual ranged from 0.177 to 0.999. Estimated abundances increased from 150 to 264 pre-flood (primary periods 1 and 2), decreased to 53 after the flood (primary period 3), an 80% decrease, and increased to 133 in primary period 4. Apparent survival was 0.985 pre-flood (interval 1), decreased to 0.100 during the flood (interval 2), and increased to 0.919 after the flood (interval 3).

At the lower site, shellfish tags were affixed to 117 *C. petrina*, 300 *C. pustulosa*, and 405 *A. plicata* among the five primary periods. For *C. petrina*, the probability of detecting a shellfish-tagged individual ranged from 0.084 to 0.999 (Figure 3.5; Appendix 3.3 & 3.4). Estimated abundances decreased from 77 to 69 pre-flood (primary periods 1 and 2), decreased to 20 after the flood (primary period 3), a 71% decrease, and decreased to zero between primary periods 4 and 5. Differences were not detected in apparent survival between intervals. For *C. pustulosa*, the probability of detecting a shellfish-tagged individual ranged from 0.087 to 1.0. Estimated abundances decreased from 202 to 163 pre-flood (primary periods 1 and 2), decreased to 68 after the flood (primary period

3), a 58% loss, and increased to 167 between primary periods 3 and 4, then decreased to 57 between primary periods 4 and 5. Apparent survival was 0.812 pre-flood (interval 1), decreased to 0.145 during the flood (interval 2), and increased to 0.844 after the flood (interval 3). For *A. plicata*, the probability of detecting a shellfish-tagged individual ranged from 0.135 to 1.0. Estimated abundances decreased from 283 to 97 pre-flood (primary periods 1 and 2); differences in estimated abundances were not detected after the flood. Differences were not detected in apparent survival between intervals.

A total of 56 pairwise comparisons were conducted to compare detection probability, estimated abundances, and apparent survival between the two data types (Appendix 3.5). Six (30%) of 20 comparisons of detection probability differed significantly; shellfish tag detection probabilities were underestimated in five (83%) of those six comparisons relative to PIT tags. Thirteen (65%) of the 20 estimated abundance comparisons differed; estimated abundances from shellfish tag data were underestimated in 10 (77%) of those 13 comparisons. Apparent survival estimates differed significantly between PIT tags and shellfish tags in one (6%) comparison; the shellfish tag estimate was lower than the PIT tag estimate in that one case. In 16 out of the 17 cases in which parameter estimates differed between tag types, there was also a change in the trend from one time period to the next (i.e. increase or decrease), indicating that these differences have the potential to influence the results of monitoring studies.

Discussion

Estimated abundances and apparent survival were directly quantified for four species of mussels during base flow and before and after large flood events, while accounting for imperfect detection. Our expectations that large floods would result in

reductions of mussel abundances and decreases in apparent survival were partially supported. Estimated abundances, but not apparent survival, of PIT-tagged *C. petrina* and *C. pustulosa* were reduced with the lesser flood magnitude (i.e., 1,283 m^3/s), and estimated abundances and apparent survival of PIT-tagged *C. petrina* and *C. pustulosa* were reduced with the greater flood magnitude (i.e., 4,332 m^3/s). However, reductions in estimated abundances of mussels with shellfish tags were not detected for two (*T. verrucosa* at the upper site and *A. plicata* at the lower site*)* of the four species, and reductions in apparent survival were not detected for three (*C. petrina, T. verrucosa* and *A. plicata*) of the four species. Based on model comparisons, estimated abundances using shellfish tags generally underestimated abundances and were more variable relative to PIT tags, which might explain the lack of concordance in model estimates between PIT-tagged *C. petrina* and *C. pustulosa* (decreased in estimated abundances with floods) and *T. verrucosa* and *A. plicata* (no detectable effects in estimated abundances with floods). Alternatively, *T. verrucosa* and *A. plicata* might be more resistant to downstream displacement during flood events than *C. petrina* or *C. pustulosa*.

At both sites, the four mussel species persisted, yet abundances decreased ranging from 0 to 80% following two large floods (1,283 m^3/s and 4,332 m^3/s), which, to date, are the highest flows recorded in a study assessing influence of large floods on mussel populations. Combining estimated abundance decreases of different mussel species with two other studies (Inoue et al., 2014; Sansom et al., 2018) for a total of five flood events within drainage basins of different sizes and habitat types, mussel responses have an apparent non-linear relationship with peak flow, ranging between 25 m^3/s and 4,332 m^3/s. Abundances of mussels were unaffected by smaller peak flows (25 to 470 m^3/s; Inoue *et*

al., 2014; Sansom *et al.*, 2018) and reported herein to decrease from 0 to 45% among three species at 1,283 m^3/s and 0 to 80% among three species at 4,332 m^3/s. However, there are several confounding factors that, once quantified and replicated, could add robustness to explaining mussel abundance-peak flow relationships. Confounding factors include site-specific attributes, such as basin size, stream geomorphology, substrate type and stability, and habitat types, and species-specific attributes such as shell morphology and burrowing behavior (Strayer, 1999; Morales *et al.*, 2006; Gangloff & Feminella, 2007; Zigler *et al.*, 2008; Allen & Vaughn, 2010; Meador *et al.*, 2011; Randklev *et al.*, 2019). Observations during this study provide empirical support for two of the reported confounding factors and their influence of mussel-abundance flow relationships. Substrates at the upper site (i.e., sand, gravel, and cobble substrates) appeared to be less scoured than the substrates at the lower site (i.e., sand and cemented sandstone substrates), suggesting a substrate-stability influence on mussel displacement. In addition, *T. verrucosa* and *A. plicata* have medial sculptured shells, which are thought to enhance anchoring ability compared to other shell sculpture types (e.g. *C. petrina*) and unsculptured species (typical central Texas form of *C. pustulosa*; Watters, 1994; Allen & Vaughn, 2009; Hornbach, Kurth & Hove, 2010; Howells, 2014; Goodding *et al.*, 2019), although the influence of shell morphology on dislodgment resistance needs further assessment (Levine, Hansen, & Gerald, 2014). Thus, targeted investigation into the relationship between peak flood discharge in different systems (e.g. small to large order rivers), habitats, substrate types, and among different mussel species is warranted.

Notable limitations of this study, and other mark-recapture studies, are the unknown fate of the mussels displaced during flood events and the lack of power to infer

mussel responses at the reach scale. Mussels displaced downstream have the potential to survive and establish new mussel beds (Hastie et al., 2001), assuming they are deposited in suitable habitat. If deposited in non-suitable habitat, the fate is less certain given their low vagility. As for reach scale inference, mussel responses to floods are heterogenous within a reach. In a 3-km reach of the Delaware River, USA, mussel aggregations were less persistent in areas with scour and compared to areas with minimal scour following multiple floods (Maloney, Lellis, Bennett, & Waddle, 2012). Therefore, 71% reductions in *C. petrina* and 80% reduction in *C. pustulosa* at the lower site with observed scour are likely overestimates of displaced *C. petrina* and *C. pustulosa* with respect to the entire lower Colorado River. This is also supported by field surveys taken in the lower Colorado River during the same period, although the surveys were not specifically designed to assess effects of the Hurricane Harvey flood on the mussel community. Prior to the flood, Ruppel (2019) reported 10.4 mussels per habitat surveyed (N = 179) with species relative abundances of 16% for *C. pustulosa* and 1.4% for *C. petrina* from a total of 1,859 mussels collected. After the flood, 6.2 mussels per habitat (N = 66) were reported with species relative abundances of 21% for *C. pustulosa* and 0.9% for *C. petrina* from a total of 441 mussels collected. Although comparing the community effects pre-flood (March – August 2017) and after the flood (September & October 2017) has limitations (e.g., unequal sampling effort, different seasons, taken at different sites within the reach), numbers of mussels per habitat and community structure suggest that reach scale reductions were less than those reported from the mark-recapture site. In future studies, establishing a patchwork of mark-recapture sites within areas with various levels of scouring potential (e.g., numerous hydraulic and substrate types) would enable greater

inferences into reach-scale effects.

With PIT tags and shellfish tags affixed to *C. petrina* and *C. pustulosa*, pairwise comparisons of model estimates using the closed robust design demonstrated that shellfish tags generally underestimated population parameters, as previously reported (Kurth, Loftin, Zydlewski, & Rhymer, 2007; Strayer & Smith, 2003; Wisniewski et al., 2013). For example, the decrease in estimated abundance of *C. petrina* at the upper sites was 31% with PIT tags, but a decrease in estimated abundance was not detected with shellfish tags. However, the decrease in estimated abundance of *C. petrina* at the lower site was 73% with PIT tags and 71% with shellfish tags. In addition to a lack of concordance in estimating abundance, detection probabilities using shellfish tags differed 75% of the time from PIT tags and were lower and more variable. This is corroborated by a previous study which found higher recapture rates for PIT tagged mussels (72 - 80%) relative to shellfish tagged mussels (30 - 47%; Kurth *et al.*, 2007). It should be noted that PIT tags have the potential to overestimate parameters relative to shellfish tags due to a higher probability of tag loss, but our analysis of PIT tag data included an estimation of tag loss. One potential mechanism for the discrepancy between shellfish and PIT tags is vertical migration through the substrate. Mussels migrate vertically for a variety of reasons, including reproduction, feeding, and thermoregulation, thus influencing the likelihood of capture with visual and tactile surveys to a large degree, and PIT tag surveys to a lesser extent because of their ability to detect subsurface mussels (Amyot & Downing, 1997; G. Thomas Watters, O'Dee, & Chordas, 2001). However, despite the differences in estimates between the two tagging methods found in this study, the general patterns in the estimates were similar and reflect the effect of floods on mussel

populations.

Studies geared towards resistance and resiliency of aquatic organisms to flooding are partly driven by the distinctly human perception that floods are devasting (i.e., considered natural disasters, destruction of human life and property). However, a differing ecological perspective has been forming through time based on the tenets of the Flow Pulse Concept (Junk, Bayley, & Sparks, 1989) and the Natural Flow Paradigm (Poff et al., 1997). Where floods are a component of the natural flow regime and not exacerbated by anthropogenic alterations (Konrad, 2003), labeling of floods as a conservation concern due to apparent localized mortality of mussels (Hastie *et al.*, 2001) might overlook long-term ecosystem services and functions of floods. Neither of the floods documented in this study were the highest flow peaks measured by USGS since 1915. Previous to 1915 and extending back to the beginning of the Holocene, flow magnitudes in western gulf slope drainages of Texas were estimated to be four to eight times greater than current magnitudes (Baker & Penteado-Orellana, 1977; Sylvia & Galloway, 2006) and well within the likely timeframe of current species radiation within the Colorado River (Inoue, Harris, Robertson, Johnson, & Randklev, 2019). Therefore, contemporary floods might not be a threat to the long-term viability of mussel populations given that this study reveals that mussels have some level of resistance to displacement despite previous findings of 4 to 8% mortalities (Hastie et al., 2001). More documentation is needed to quantify short-term mussel resistance (e.g., site-level and reach-level responses to floods while accounting for confounding factors), but even more important, quantification of mussel population resiliency is needed to fully understand ecosystem services and functions and of large floods on the long-term fitness of mussels.

Table 3.1. Discharge values (m^3/s) for each sampling site for each interval (i.e., the time between primary periods; the first interval occurred between primary period one and primary period two). Included are median discharge, minimum discharge, and maximum discharge.

Site	Interval	Minimum	Median	Maximum
Upper	1	0.60	1.10	2.10
	2	1.11	1.81	76.46
	3	0.09	0.79	31.89
	4	0.06	11.02	1,270.29
Lower	1	21.58	33.61	52.25
	2	19.03	29.34	4,336.50
	3	12.97	17.36	618.89
	4	22.63	32.85	87.16

Table 3.2. The total number of mussels (Total N) sampled at each site with survey gear for each species.

Site	Gear	Species	Total N
Upper	PIT tags	*C. petrina*	442
		C. pustulosa	15
	Shellfish tags	*C. petrina*	442
		C. pustulosa	15
		T. verrucosa	217
Lower	PIT tags	*C. petrina*	120
		C. pustulosa	308
	Shellfish tags	*C. petrina*	117
		C. pustulosa	300
		A. plicata	405

Table 3.3. Summary of estimated abundance and apparent survival of four mussel species after flood events at the upper and lower sites in the Colorado River. Estimated abundance percentage indicates the percent decrease in estimated abundance between the primary periods surrounding the flood for those comparisons which were significant. The letters "NS" represent no significant difference detected between primary periods (estimated abundance) or intervals (apparent survival). The symbol "--" represents species that did not have PIT tags.

		PIT-tagged		Shellfish-tagged	
Site	Species	Estimated abundance	Apparent survival	Estimated abundance	Apparent survival
Upper site	*C. petrina*	decreased 31%	NS	NS	NS
	C. pustulosa	decreased 45%	NS	NS	NS
	T. verrucosa	--	--	NS	NS
Lower site	*C. petrina*	decreased 73%	decreased	decreased 71%	NS
	C. pustulosa	decreased 80%	decreased	decreased 58%	decreased
	A. plicata	--	--	NS	NS

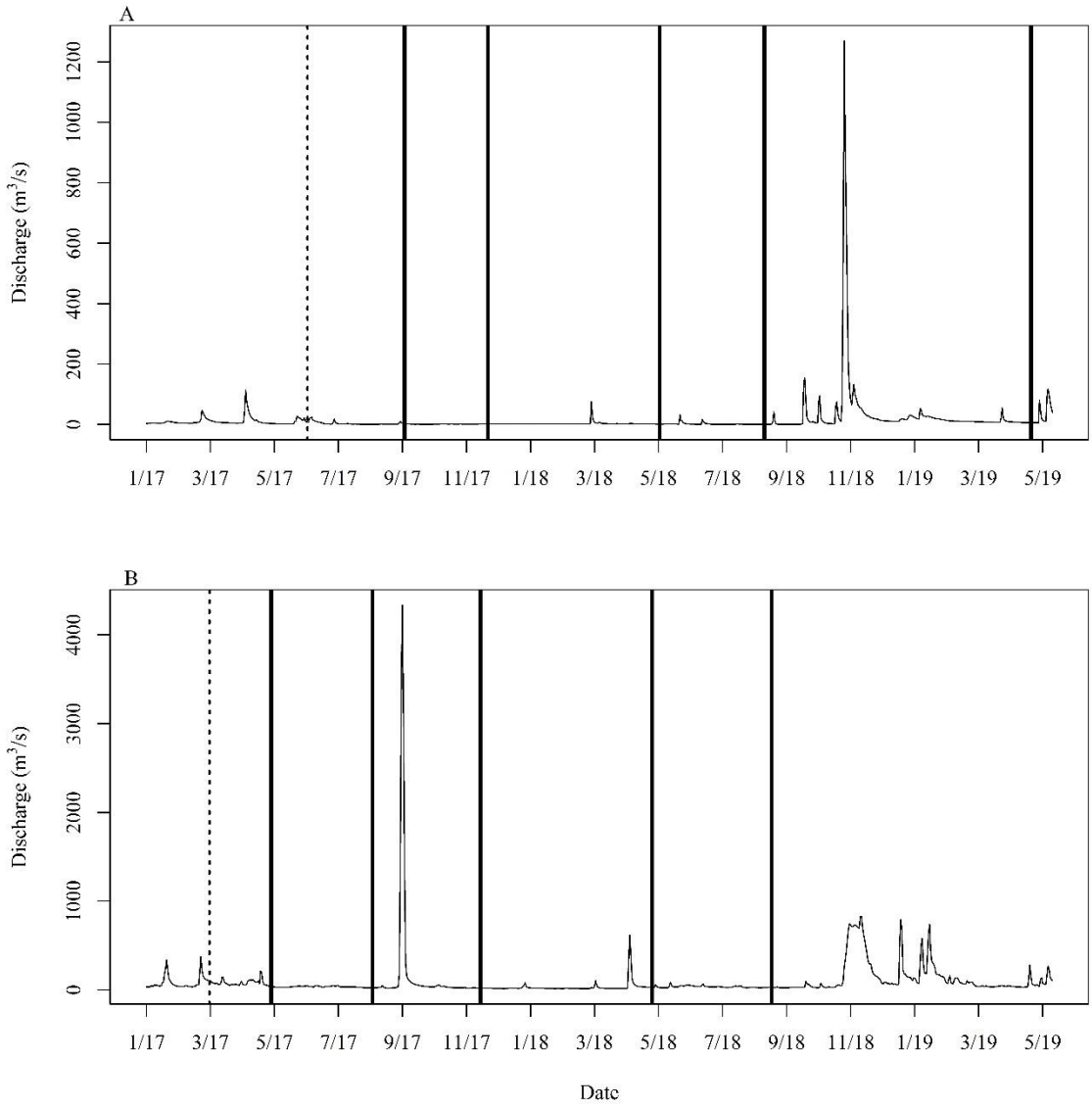


Figure 3.1. Discharge (m^3/s) plot of the upper (A) and lower (B) Colorado River throughout the duration of the study periods taken from USGS gages 08147000 (upper site) and 08161000 (lower site). Black dotted line denotes initial tagging event, black solid lines denote primary period sampling events.

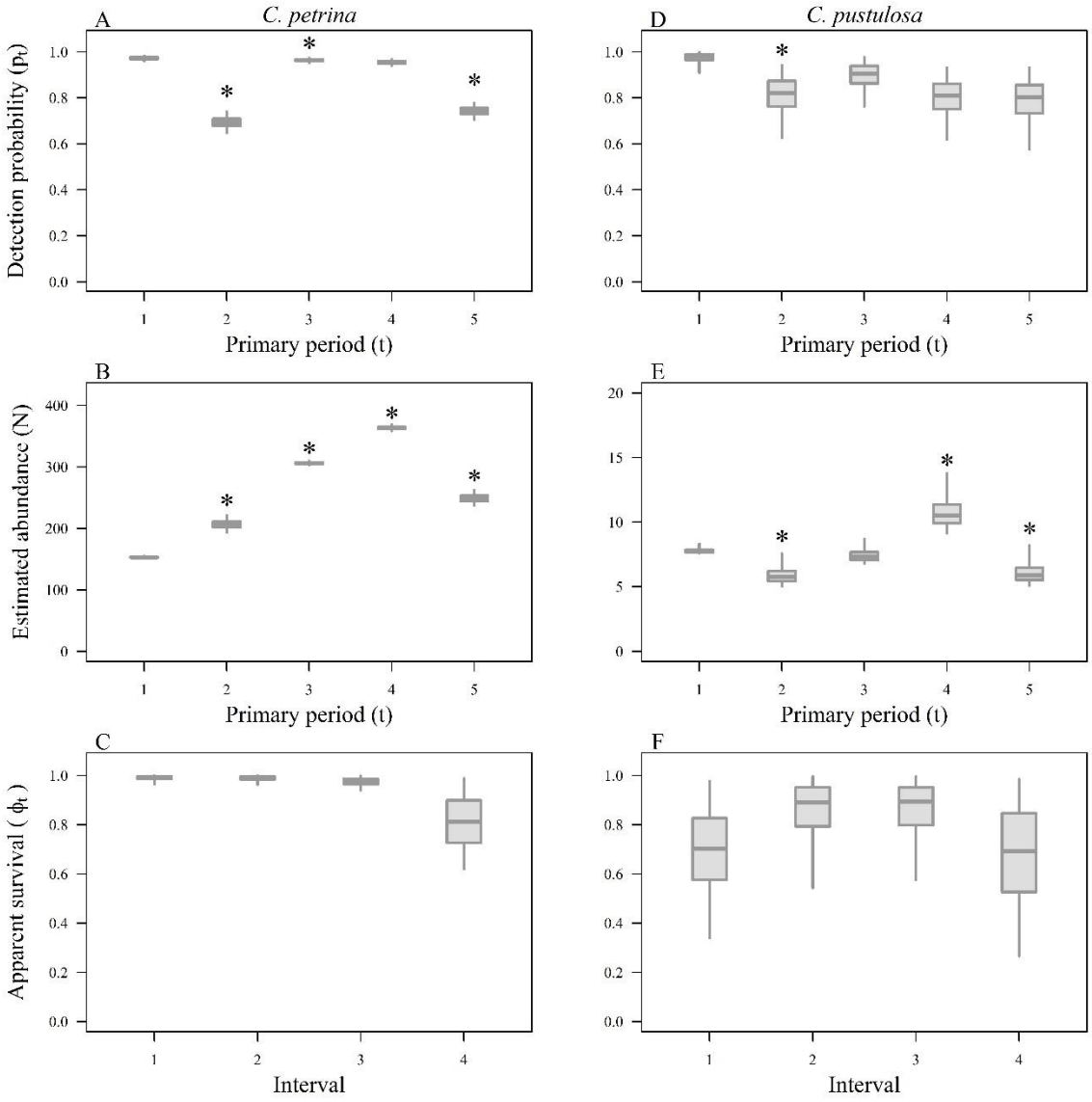


Figure 3.2. Boxplots depicting mark-recapture estimates of detection probability, estimated abundance, and apparent survival for *C. petrina* (A-C) and *C. pustulosa* (D-F) from the upper site on the Colorado River using PIT tag data. The flood occurred between primary period 4 and 5 and interval 4. An * indicates a significant difference between the primary period or interval and the one directly preceding it in time (i.e. *t-1* to *t*).

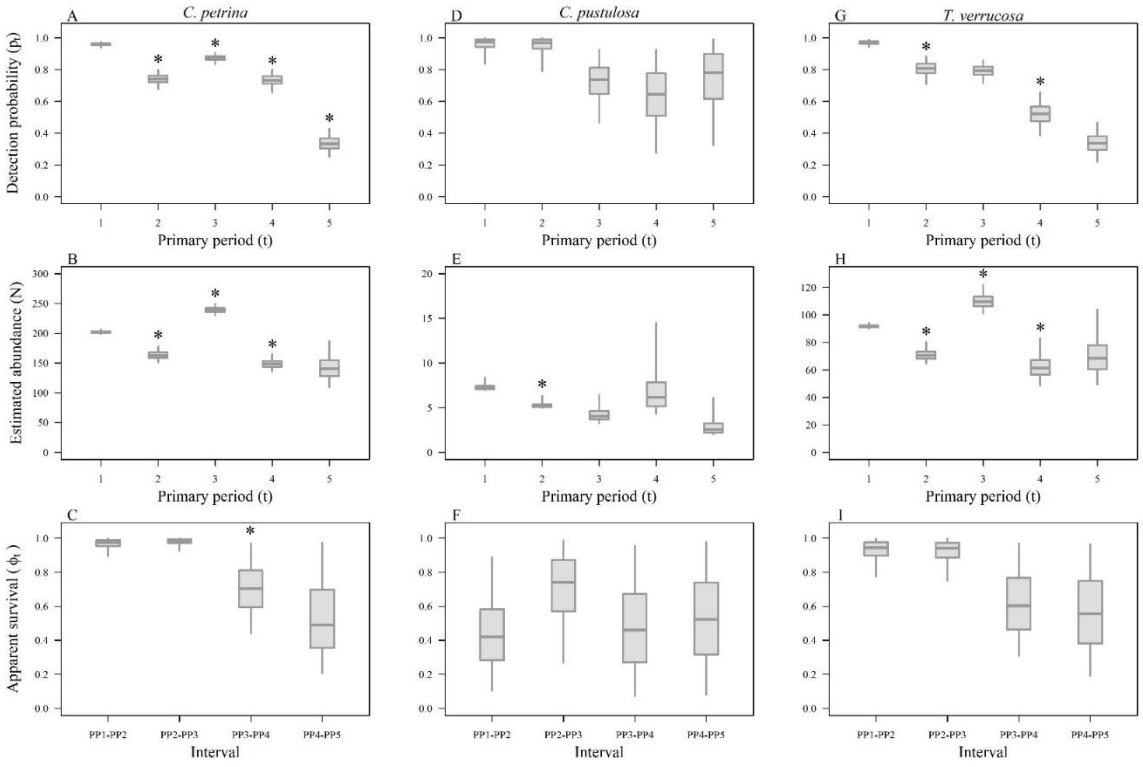


Figure 3.3. Boxplots depicting mark-recapture estimates of detection probability, abundance, and apparent survival for *C. petrina* (A-C), *C. pustulosa* (D-F), and *T. verrucosa* (G-I) from the upper site on the Colorado River using shellfish tag data. The flood occurred between primary period 4 and 5 and interval 4. An * indicates a significant difference between the primary period or interval and the one directly preceding it in time (i.e. *t-1* to *t*).

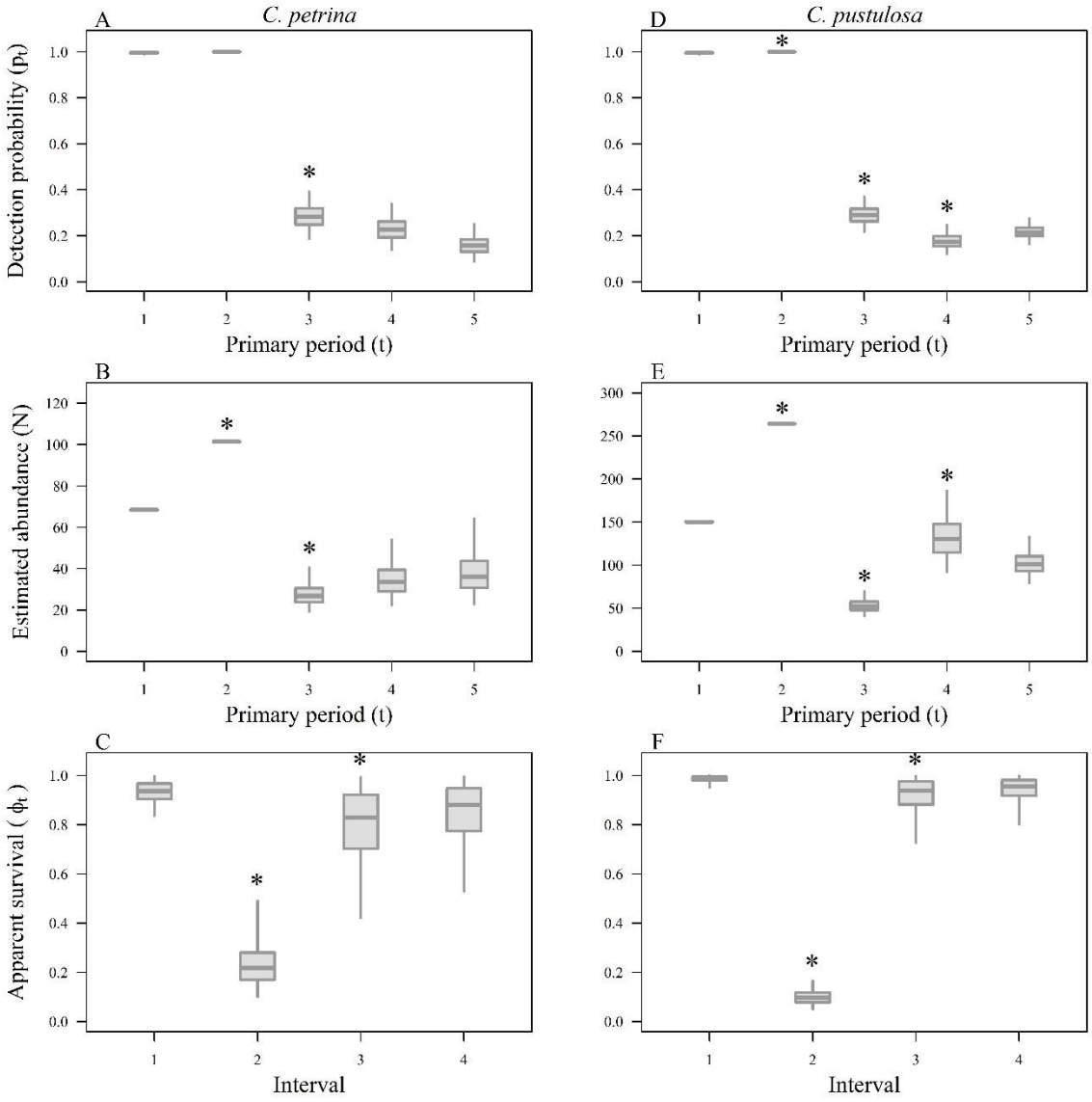


Figure 3.4. Boxplots depicting mark-recapture estimates of detection probability, abundance, and apparent survival for *C. petrina* (A-C) and *C. pustulosa* (D-F) from the lower site on the Colorado River using PIT tag data. The flood occurred between primary period 2 and 3 and interval 2. An * indicates a significant difference between the primary period or interval and the one directly preceding it in time (i.e. *t-1* to *t*).

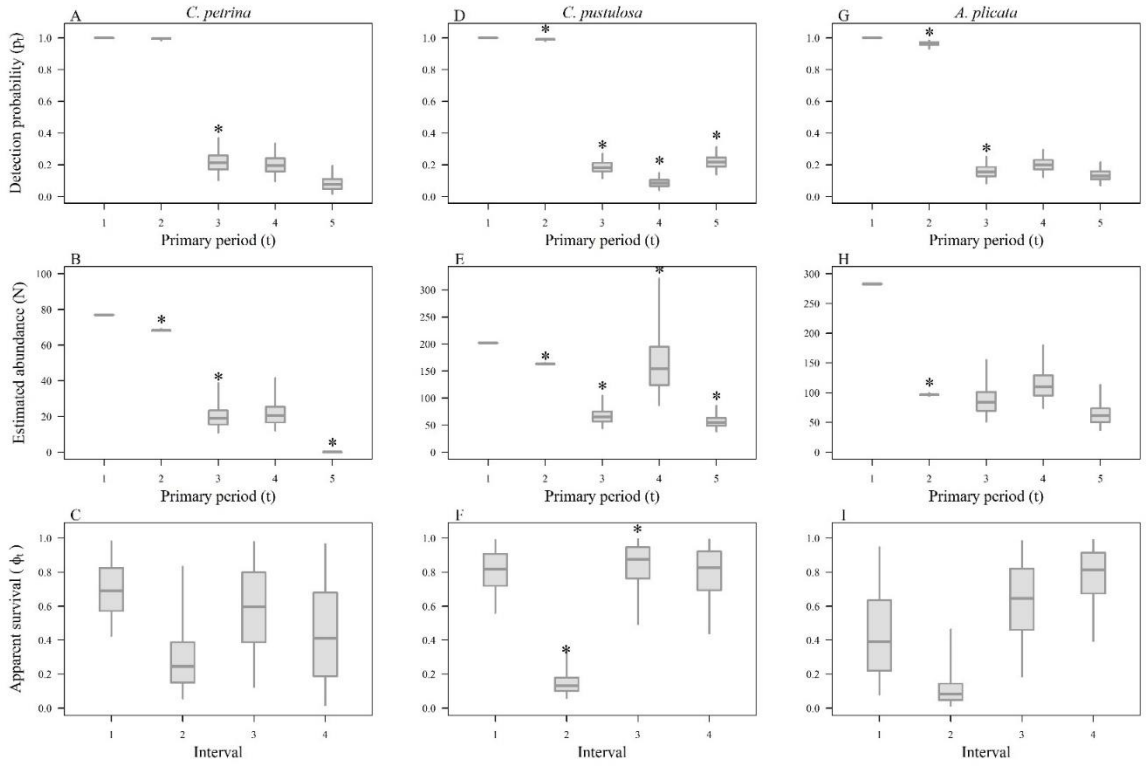


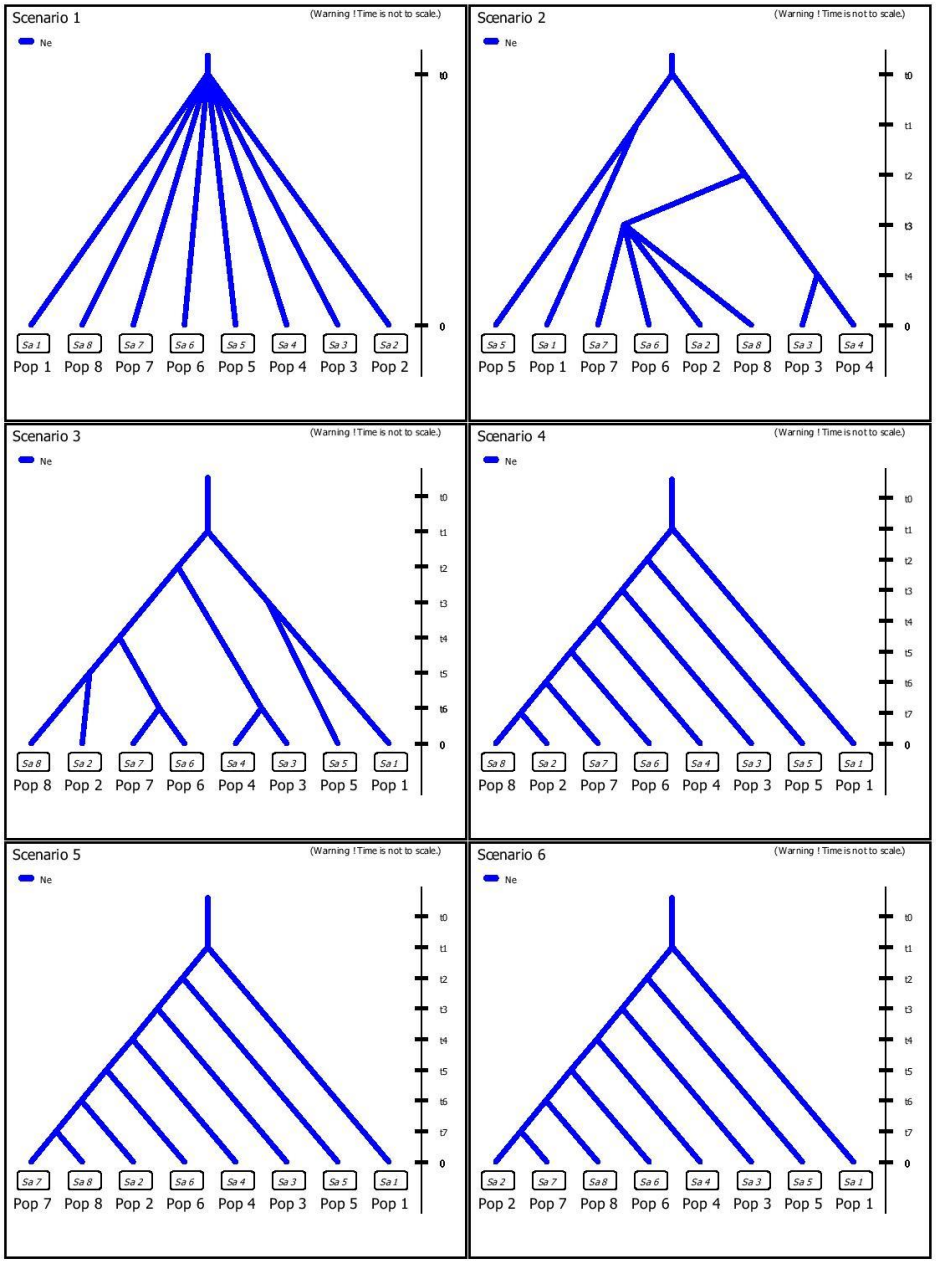
Figure 3.5. Boxplots depicting mark-recapture estimates of detection probability, abundance, and apparent survival for *C. petrina* (A-C), *C. pustulosa* (D-F), and *A. plicata* (G-I) from the lower site on the Colorado River using shellfish tag data. The flood occurred between primary period 2 and 3 and interval 2. An * indicates a significant difference between the primary period or interval and the one directly preceding it in time (i.e. *t-1* to *t*)

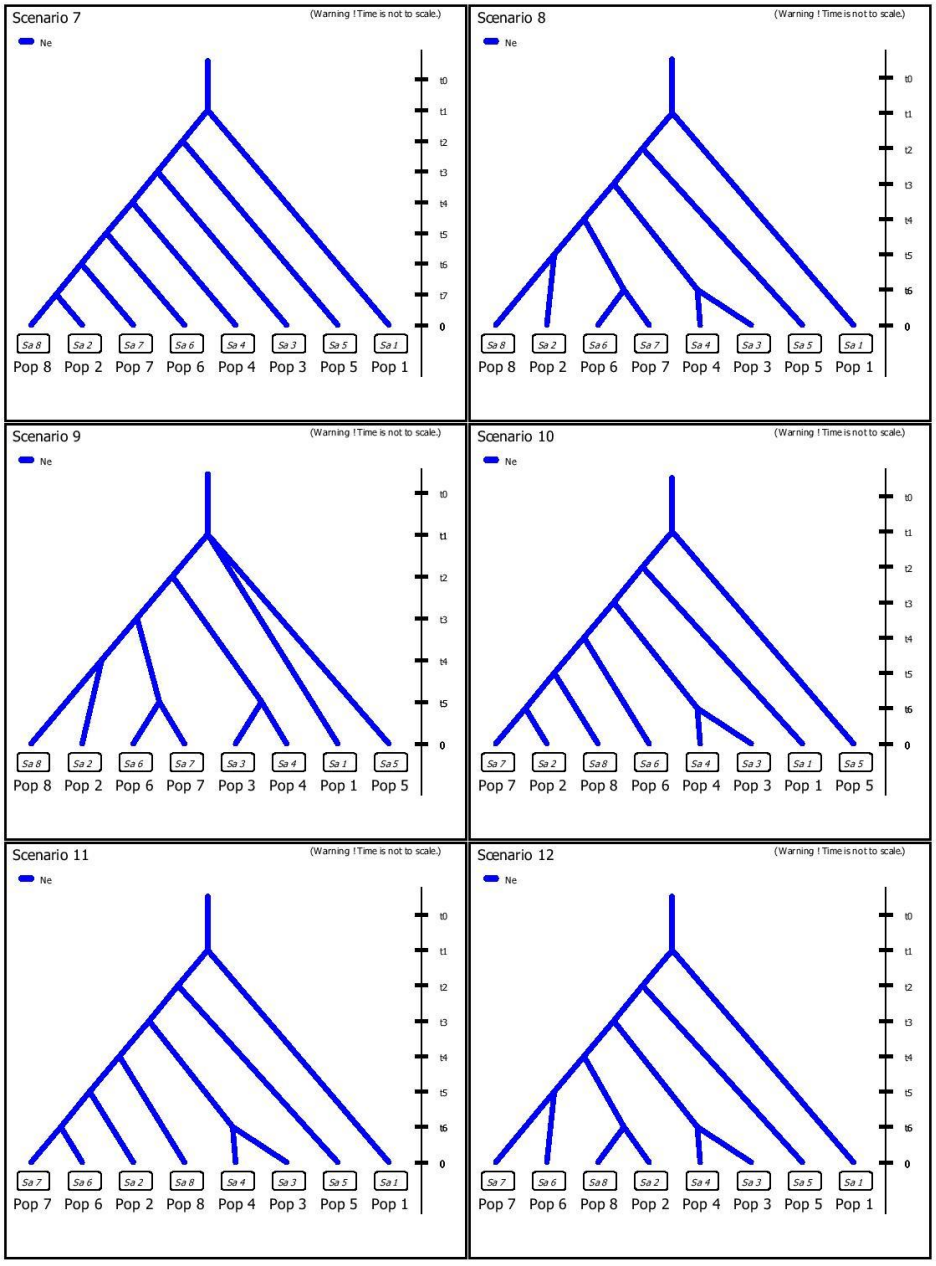
APPENDIX SECTION

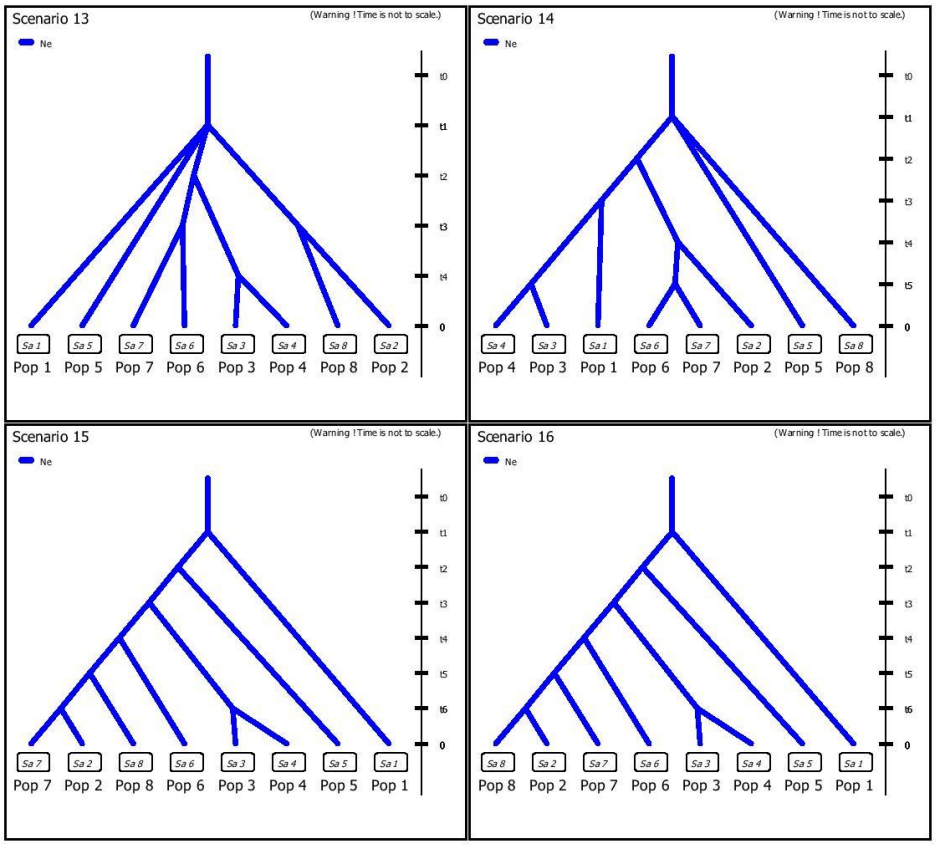
LIST OF APPENDICES

Appendix **Page**

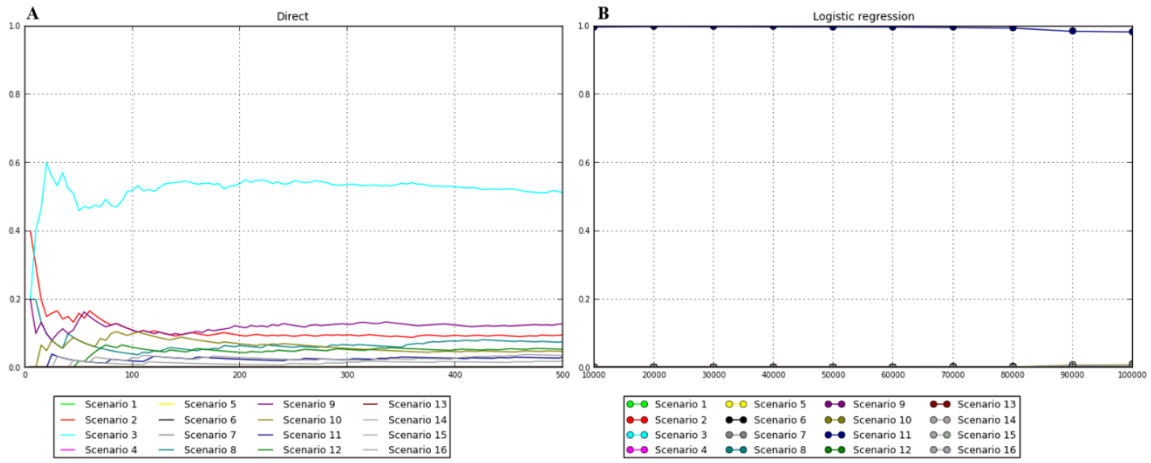
Appendix 2.1. DIYABC scenarios tested. See Table 2.3 for details and sources. Briefly, scenario 1 is the null model of a simultaneous model of divergence, scenario 2 was designed from phylogenetic topology, scenario 3 was designed from Entropy plots, scenario 4-8 are various models of peripheral isolation or vicariance from Eisenhour (2004), scenario 9 is based off of G_{ST} relationships, scenario 10 is modeled off of Underwood et al. (2003), scenario 11, 12, 15, and 16 are based off of stepping-stone model of dispersal, scenario 13 and 14 is based off Echelle et al. (2018). Pop 1- *M. aestivalis*, Pop 2- *M. australis*, Pop 3- BrCol *M. hyostoma*, Pop 4- C.F. *M. hyostoma*, Pop 5- *M. marconis*, Pop 6- M.R. *M. hyostoma*, Pop 7- R.R. *M. hyostoma*, Pop 8- *M. tetranema*.




Scenario 7
(Warning ! Time is not to scale.)
Ne
t0 t1 t2 t3 t4 t5 t6 t7 0
Sa 8 Sa 2 Sa 7 Sa 6 Sa 4 Sa 3 Sa 5 Sa 1
Pop 8 Pop 2 Pop 7 Pop 6 Pop 4 Pop 3 Pop 5 Pop 1
Scenario 8
(Warning ! Time is not to scale.)
Ne
t0 t1 t2 t3 t4 t5 t6 0
Sa 8 Sa 2 Sa 6 Sa 7 Sa 4 Sa 3 Sa 5 Sa 1
Pop 8 Pop 2 Pop 6 Pop 7 Pop 4 Pop 3 Pop 5 Pop 1
Scenario 9
(Warning ! Time is not to scale.)
Ne
t0 t1 t2 t3 t4 t5 0
Sa 8 Sa 2 Sa 6 Sa 7 Sa 3 Sa 4 Sa 1 Sa 5
Pop 8 Pop 2 Pop 6 Pop 7 Pop 3 Pop 4 Pop 1 Pop 5
Scenario 10
(Warning ! Time is not to scale.)
Ne
t0 t1 t2 t3 t4 t5 t6 0
Sa 7 Sa 2 Sa 8 Sa 6 Sa 4 Sa 3 Sa 1 Sa 5
Pop 7 Pop 2 Pop 8 Pop 6 Pop 4 Pop 3 Pop 1 Pop 5
Scenario 11
(Warning ! Time is not to scale.)
Ne
t0 t1 t2 t3 t4 t5 t6 0
Sa 7 Sa 6 Sa 2 Sa 8 Sa 4 Sa 3 Sa 5 Sa 1
Pop 7 Pop 6 Pop 2 Pop 8 Pop 4 Pop 3 Pop 5 Pop 1
Scenario 12
(Warning ! Time is not to scale.)
Ne
t0 t1 t2 t3 t4 t5 t6 0
Sa 7 Sa 6 Sa 8 Sa 2 Sa 4 Sa 3 Sa 5 Sa 1
Pop 7 Pop 6 Pop 8 Pop 2 Pop 4 Pop 3 Pop 5 Pop 1


Scenario 13
(Warning ! Time is not to scale.)
Ne
t0
t1
t2
t3
t4
0
Sa 1
Sa 5
Sa 7
Sa 6
Sa 3
Sa 4
Sa 8
Sa 2
Pop 1
Pop 5
Pop 7
Pop 6
Pop 3
Pop 4
Pop 8
Pop 2
Scenario 14
(Warning ! Time is not to scale.)
Ne
t0
t1
t2
t3
t4
t5
0
Sa 4
Sa 3
Sa 1
Sa 6
Sa 7
Sa 2
Sa 5
Sa 8
Pop 4
Pop 3
Pop 1
Pop 6
Pop 7
Pop 2
Pop 5
Pop 8
Scenario 15
(Warning ! Time is not to scale.)
Ne
t0
t1
t2
t3
t4
t5
t6
0
Sa 7
Sa 2
Sa 8
Sa 6
Sa 3
Sa 4
Sa 5
Sa 1
Pop 7
Pop 2
Pop 8
Pop 6
Pop 3
Pop 4
Pop 5
Pop 1
Scenario 16
(Warning ! Time is not to scale.)
Ne
t0
t1
t2
t3
t4
t5
t6
0
Sa 8
Sa 2
Sa 7
Sa 6
Sa 3
Sa 4
Sa 5
Sa 1
Pop 8
Pop 2
Pop 7
Pop 6
Pop 3
Pop 4
Pop 5
Pop 1

Appendix 2.2. Plots of direct (A) and logistic (B) probabilities of each scenario tested in DIYABC.



Appendix 3.1. Mark-recapture median estimate, lower bound (LB), and upper bound (UB) credible intervals of detection probability and estimated abundances for each primary period (e.g. PP1 is primary period 1), and apparent survival for each interval for *C. petrina* and *C. pustulosa* from the upper and lower Colorado River via PIT surveys.

Site	Species	Parameter	Period	Estimate	LB	UB
Upper	*C. petrina*	Detection Probability	PP1	0.972	0.959	0.983
			PP2	0.693	0.647	0.739
			PP3	0.963	0.951	0.973
			PP4	0.954	0.94	0.967
			PP5	0.743	0.705	0.778
		Estimated Abundances	PP1	152.937	151.3	155.01
			PP2	206.563	193.378	220.87
			PP3	305.792	302.74	309.592
			PP4	363.252	358.331	368.917
			PP5	248.833	237.321	261.929
		Apparent Survival	Interval 1	0.99	0.965	1.0
			Interval 2	0.989	0.962	1.0
			Interval 3	0.975	0.94	0.998
			Interval 4	0.811	0.62	0.988
	C. pustulosa	Detection Probability	PP1	0.973	0.91	0.998
			PP2	0.812	0.626	0.942
			PP3	0.895	0.762	0.976
			PP4	0.8	0.619	0.931
			PP5	0.789	0.577	0.932
		Estimated Abundances	PP1	7.789	7.591	8.325
			PP2	5.894	5.028	7.565
			PP3	7.436	6.789	8.696
			PP4	10.78	9.155	13.775
			PP5	6.094	5.082	8.21
		Apparent Survival	Interval 1	0.695	0.34	0.976
			Interval 2	0.857	0.543	0.996
			Interval 3	0.862	0.576	0.996
			Interval 4	0.678	0.267	0.985
Lower	*C. petrina*	Detection Probability	PP1	0.996	0.989	0.999
			PP2	0.999	0.998	1.0
			PP3	0.284	0.187	0.392
			PP4	0.229	0.14	0.34
			PP5	0.16	0.089	0.249
		Estimated Abundances	PP1	68.461	68.239	68.916
			PP2	101.408	101.342	101.554
			PP3	27.642	19.338	40.556
			PP4	34.869	22.314	54.086
			PP5	38.105	22.8	64.034
		Apparent Survival	Interval 1	0.933	0.837	0.997
			Interval 2	0.237	0.1	0.49
			Interval 3	0.796	0.422	0.993
			Interval 4	0.845	0.53	0.995
	C. pustulosa	Detection Probability	PP1	0.995	0.99	0.998
			PP2	0.999	0.999	1.0
			PP3	0.29	0.218	0.37
			PP4	0.177	0.122	0.247
			PP5	0.216	0.164	0.275
		Estimated Abundances	PP1	150.334	149.871	151.16
			PP2	264.389	264.277	264.575

	PP3	53.181	40.965	69.47
	PP4	132.787	92.102	186.389
	PP5	102.456	79.339	132.813
Apparent Survival	Interval 1	0.985	0.951	1.0
	Interval 2	0.1	0.051	0.164
	Interval 3	0.919	0.727	0.998
	Interval 4	0.942	0.802	0.999

Appendix 3.2. Median differences and lower bound (LB) and upper bound (UB) credible intervals between periods for each site, species, and parameter. Period column indicates which primary periods (e.g., 1-2 indicates PP1-PP2 comparison) or intervals (e.g., 1-2 indicates first and second interval comparison) were compared. An * in the period column indicates a significant difference between the respective periods or intervals. Data in this table are for PIT tag estimates.

Site	Species	Parameter	Periods	Median Difference	LB	UB
Upper	*C. petrina*	Detection Probability	1-2*	0.279	0.231	0.327
			2-3*	-0.270	-0.318	-0.223
			3-4	0.009	-0.009	0.026
			4-5	0.211	0.174	0.252
		Estimated Abundance	1-2*	-53.410	-68.158	-40.299
			2-3*	-99.511	-112.744	-84.314
			3-4*	-57.337	-63.831	-51.239
			4-5*	114.553	100.266	126.908
		Apparent Survival	1-2	0.001	-0.028	0.030
			2-3	0.013	-0.022	0.053
			3-4	0.162	-0.018	0.359
	C. pustulosa	Detection Probability	1-2*	0.152	0.016	0.351
			2-3	-0.079	-0.293	0.115
			3-4	0.090	-0.095	0.303
			4-5	0.008	-0.235	0.265
		Estimated Abundance	1-2*	2.021	0.190	2.924
			2-3	-1.590	-3.187	0.263
			3-4*	-3.132	-6.454	-1.237
			4-5*	4.574	1.923	8.071
		Apparent Survival	1-2	-0.171	-0.565	0.281
			2-3	-0.003	-0.363	0.336
			3-4	0.172	-0.237	0.649
Lower	*C. petrina*	Detection Probability	1-2	-0.003	-0.010	0.000
			2-3*	0.717	0.606	0.812
			3-4	0.056	-0.087	0.197
			4-5	0.069	-0.056	0.198
		Estimated Abundance	1-2*	-32.986	-33.219	-32.481
			2-3*	74.547	60.863	82.115
			3-4	-6.515	-28.290	10.295
			4-5	-2.502	-32.057	22.095
		Apparent Survival	1-2*	0.713	0.433	0.859
			2-3*	-0.608	-0.849	-0.033
			3-4	-0.043	-0.473	0.358
	C. pustulosa	Detection Probability	1-2*	-0.004	-0.009	-0.001
			2-3*	0.710	0.630	0.781
			3-4*	0.114	0.013	0.208
			4-5	-0.040	-0.119	0.049
		Estimated Abundance	1-2*	-114.105	-114.567	-113.229
			2-3*	212.001	194.989	223.445
			3-4*	-77.466	-133.207	-35.925
			4-5	28.559	-21.258	89.693
		Apparent Survival	1-2*	0.888	0.811	0.939
			2-3*	-0.839	-0.931	-0.609
			3-4	-0.013	-0.235	0.152

Appendix 3.3. Mark-recapture median estimate, lower bound (LB), and upper bound (UB) credible intervals of detection probability and estimated abundances for each primary period (e.g. PP1 is primary period 1), and apparent survival for each interval for *C. petrina* (upper and lower sites), *C. pustulosa* (upper and lower sites), *T. verrucosa* (upper site), and *A. plicata* (lower site) via tactile surveys.

Site	Species	Parameter	Period	Estimate	LB	UB
Upper	*C. petrina*	Detection Probability	PP1	0.958	0.94	0.974
			PP2	0.741	0.676	0.802
			PP3	0.872	0.832	0.906
			PP4	0.734	0.659	0.801
			PP5	0.335	0.251	0.425
		Estimated Abundances	PP1	202.178	198.914	206.193
			PP2	163.465	150.6	178.857
			PP3	239.543	230.288	250.788
			PP4	148.773	135.952	165.209
			PP5	142.671	110.488	186.714
		Apparent Survival	Interval 1	0.967	0.896	0.999
			Interval 2	0.978	0.925	0.999
			Interval 3	0.706	0.447	0.973
			Interval 4	0.53	0.208	0.97
	C. pustulosa	Detection Probability	PP1	0.958	0.832	0.999
			PP2	0.946	0.774	0.999
			PP3	0.724	0.458	0.923
			PP4	0.639	0.297	0.933
			PP5	0.755	0.347	0.991
		Estimated Abundances	PP1	7.315	6.997	8.405
			PP2	5.304	4.998	6.45
			PP3	4.281	3.247	6.538
			PP4	6.922	4.282	13.451
			PP5	2.878	2.016	5.764
		Apparent Survival	Interval 1	0.44	0.109	0.893
			Interval 2	0.712	0.288	0.986
			Interval 3	0.489	0.076	0.963
			Interval 4	0.516	0.084	0.97
	T. verrucosa	Detection Probability	PP1	0.97	0.944	0.987
			PP2	0.805	0.709	0.883
			PP3	0.791	0.714	0.859
			PP4	0.521	0.385	0.658
			PP5	0.339	0.221	0.466
		Estimated Abundances	PP1	91.72	90.119	94.229
			PP2	70.975	64.503	80.309
			PP3	110.127	101.173	121.846
			PP4	62.495	48.582	83.025
			PP5	70.346	49.314	103.847
		Apparent Survival	Interval 1	0.928	0.774	0.998
			Interval 2	0.921	0.751	0.998
			Interval 3	0.619	0.308	0.97
			Interval 4	0.566	0.191	0.965
Lower	*C. petrina*	Detection Probability	PP1	0.999	0.998	1.0
			PP2	0.995	0.984	1.0
			PP3	0.221	0.109	0.364
			PP4	0.199	0.099	0.331
			PP5	0.084	0.018	0.199
		Estimated Abundances	PP1	76.954	76.906	77.092

		PP2	68.249	67.944	69.051
		PP3	20.024	10.967	36.539
		PP4	22.111	12.051	40.173
		PP5	0.0	0.0	0.0
	Apparent Survival	Interval 1	0.683	0.423	0.972
		Interval 2	0.311	0.06	0.827
		Interval 3	0.577	0.117	0.982
		Interval 4	0.451	0.018	0.969
C. pustulosa	Detection Probability	PP1	1.0	0.999	1.0
		PP2	0.99	0.98	0.997
		PP3	0.184	0.116	0.262
		PP4	0.087	0.041	0.151
		PP5	0.219	0.144	0.313
	Estimated Abundances	PP1	201.796	201.751	201.908
		PP2	163.39	162.31	165.169
		PP3	68.093	45.694	103.405
		PP4	167.465	85.741	319.778
		PP5	56.933	38.231	83.01
	Apparent Survival	Interval 1	0.812	0.564	0.992
		Interval 2	0.145	0.059	0.284
		Interval 3	0.844	0.519	0.995
		Interval 4	0.798	0.444	0.993
A. plicata	Detection Probability	PP1	1.0	1.0	1.0
		PP2	0.964	0.933	0.984
		PP3	0.16	0.084	0.252
		PP4	0.202	0.122	0.297
		PP5	0.135	0.071	0.217
	Estimated Abundances	PP1	282.837	282.836	282.84
		PP2	96.481	94.437	99.617
		PP3	88.133	51.553	154.901
		PP4	114.331	73.908	179.96
		PP5	64.381	36.828	113.382
	Apparent Survival	Interval 1	0.437	0.079	0.947
		Interval 2	0.118	0.014	0.462
		Interval 3	0.631	0.186	0.985
		Interval 4	0.779	0.395	0.991

Appendix 3.4. Median differences and lower bound (LB) and upper bound (UB) credible intervals between periods for each site, species, and parameter. Period column indicates which primary periods (e.g., 1-2 indicates PP1-PP2 comparison) or intervals (e.g., 1-2 indicates first and second interval comparison) were compared. An * in the period column indicates a significant difference between the respective periods or intervals. Data in this table are for shellfish tag estimates.

Site	Species	Parameter	Periods	Median Diff	LB	UB
Upper	*C. petrina*	Detection Probability	1-2*	0.217	0.154	0.284
			2-3*	-0.131	-0.205	-0.058
			3-4*	0.137	0.063	0.220
			4-5*	0.400	0.283	0.505
		Estimated Abundance	1-2*	39.164	22.984	52.203
			2-3*	-76.319	-92.739	-58.018
			3-4*	91.015	72.236	107.476
			4-5	7.762	-38.887	41.997
		Apparent Survival	1-2	-0.008	-0.087	0.053
			2-3*	0.276	0.001	0.541
			3-4	0.201	-0.416	0.664
	C. pustulosa	Detection Probability	1-2	0.004	-0.134	0.195
			2-3	0.213	-0.035	0.508
			3-4	0.081	-0.314	0.496
			4-5	-0.123	-0.571	0.395
		Estimated Abundance	1-2*	2.015	0.781	3.184
			2-3	1.185	-1.321	2.637
			3-4	-1.998	-9.414	1.050
			4-5	3.448	-0.023	10.995
		Apparent Survival	1-2	-0.300	-0.776	0.383
			2-3	0.241	-0.462	0.801
			3-4	-0.036	-0.699	0.682
	T. verrucosa	Detection Probability	1-2*	0.162	0.081	0.263
			2-3	0.014	-0.100	0.129
			3-4*	0.270	0.117	0.423
			4-5	0.184	-0.001	0.363
		Estimated Abundance	1-2*	21.208	11.178	27.824
			2-3*	-38.862	-52.955	-26.525
			3-4*	48.239	25.937	65.392
			4-5	-6.683	-43.866	20.669
		Apparent Survival	1-2	0.005	-0.183	0.207
			2-3	0.314	-0.096	0.646
			3-4	0.043	-0.554	0.688
Lower	*C. petrina*	Detection Probability	1-2	0.003	-0.001	0.016
			2-3*	0.778	0.629	0.887
			3-4	0.020	-0.144	0.201
			4-5	0.114	-0.034	0.270
		Estimated Abundance	1-2*	8.786	7.895	9.038
			2-3*	49.812	31.628	57.317
			3-4	-1.805	-23.335	16.955
			4-5*	20.587	12.051	40.173
		Apparent Survival	1-2	0.388	-0.218	0.826
			2-3	-0.307	-0.837	0.538
			3-4	0.145	-0.672	0.832
	C. pustulosa	Detection Probability	1-2*	0.009	0.003	0.020
			2-3*	0.808	0.727	0.875

		3-4*	0.097	0.006	0.191
		4-5*	-0.130	-0.233	-0.034
	Estimated Abundance	1-2*	38.532	36.609	39.487
		2-3*	97.492	59.562	117.727
		3-4*	-87.360	-252.186	-11.164
		4-5*	98.349	24.964	265.558
	Apparent Survival	1-2*	0.683	0.354	0.899
		2-3*	-0.734	-0.900	-0.330
		3-4	0.037	-0.379	0.462
A. plicata	Detection Probability	1-2*	0.035	0.016	0.067
		2-3*	0.807	0.705	0.884
		3-4	-0.043	-0.162	0.082
		4-5	0.068	-0.047	0.180
	Estimated Abundance	1-2*	186.518	183.221	188.399
		2-3	12.688	-58.774	45.627
		3-4	-26.202	-100.271	50.572
		4-5	48.032	-12.427	120.247
	Apparent Survival	1-2	0.295	-0.293	0.899
		2-3	-0.553	-0.929	0.182
		3-4	-0.135	-0.697	0.388

Appendix 3.5. Median, lower bound (LB), and upper bound (UB) credible intervals of differences between parameter estimates from PIT tag and shellfish tag data for each species and both sites where a comparison was possible. If the median is negative, that indicates shellfish tag data underestimated parameter relative to PIT tag data, and overestimated if positive. An * in the period column indicates that zero is not included in the 95% credible intervals, and the posterior estimates of PIT and shellfish data are significantly different.

Site	Species	Parameter	Period	LB	Median	UB
Upper	*C. petrina*	Detection Probability	PP1	-0.036	-0.013	0.006
			PP2	-0.033	0.049	0.123
			PP3*	-0.133	-0.091	-0.055
			PP4*	-0.295	-0.219	-0.152
			PP5*	-0.497	-0.408	-0.312
		Estimated Abundance	PP1*	45.450	49.110	53.718
			PP2*	-61.583	-43.452	-22.897
			PP3*	-76.041	-66.537	-54.488
			PP4*	-228.658	-215.041	-197.705
			PP5*	-140.429	-108.053	-60.634
		Apparent Survival	PP1-PP2	-0.094	-0.270	0.022
			PP2-PP3	-0.066	-0.128	0.025
			PP3-PP4*	-0.529	-0.389	-0.001
			PP4-PP5	-0.683	-0.164	0.215
	C. pustulosa	Detection Probability	PP1	-0.141	-0.006	0.064
			PP2	-0.069	0.135	0.337
			PP3	-0.459	-0.161	0.069
			PP4	-0.538	-0.153	0.192
			PP5	-0.478	-0.009	0.302
		Estimated Abundance	PP1	-1.127	-0.552	0.602
			PP2	-2.335	-0.535	0.790
			PP3*	-4.863	-3.288	-0.718
			PP4	-8.172	-4.387	3.013
			PP5*	-5.663	-3.296	-0.119
		Apparent Survival	PP1-PP2	-0.724	-0.270	0.313
			PP2-PP3	-0.632	-0.128	0.314
			PP3-PP4	-0.840	-0.389	0.166
			PP4-PP5	-0.779	-0.164	0.490
Lower	*C. petrina*	Detection Probability	PP1	0.000	0.003	0.010
			PP2	-0.016	-0.003	0.001
			PP3	-0.219	-0.068	0.111
			PP4	-0.184	-0.032	0.138
			PP5	-0.190	-0.079	0.054
		Estimated Abundance	PP1*	8.008	8.530	8.751
			PP2*	-33.501	-33.242	-32.347
			PP3	-24.540	-8.004	11.146
			PP4	-35.951	-12.613	10.334
			PP5*	-64.034	-36.170	-22.800
		Apparent Survival	PP1-PP2	-0.532	-0.256	0.043
			PP2-PP3	-0.303	0.046	0.602
			PP3-PP4	-0.773	-0.208	0.357
			PP4-PP5	-0.910	-0.422	0.204

C. pustulosa	Detection Probability	PP1*	0.001	0.004	0.010
		PP2*	-0.020	-0.008	-0.003
		PP3	-0.217	-0.107	0.001
		PP4*	-0.175	-0.090	-0.006
		PP5	-0.096	0.002	0.108
	Estimated Abundance	PP1*	50.626	51.522	51.937
		PP2*	-102.078	-101.123	-99.225
		PP3	-13.336	13.205	53.544
		PP4	-65.785	23.641	191.430
		PP5*	-81.700	-45.631	-9.915
	Apparent Survival	PP1-PP2	-0.423	-0.160	0.009
		PP2-PP3	-0.066	0.038	0.190
		PP3-PP4	-0.425	-0.053	0.186
		PP4-PP5	-0.522	-0.114	0.104

Appendix 3.6. Model code used in R for closed robust design to estimate detection probability, apparent survival, and estimated abundance while accounting for tag loss. Code is modified from Riecke et al. 2018.

```
# Model specification for mussel mark-recapture study in the Colorado River, Texas----
modelString = "
model {
# Priors
# Survival and gamma
   for(t in 1:(n.years-1)){
     phi[t] ~ dbeta(1,1)
     gamma[t] ~ dbeta(1,1)
  }
# Secondary occasions p
   for (t in 1:n.years){
   for (j in 1:(n.sec[t])){
     p[t,j] ~ dunif(0,1)
     yes[t,j] ~ dbin(p[t,j], total[t,j])
   }
  }
#mean.p ~ dunif(0,1)
# Primary occasions p
   for (t in 1:n.years){
     pstar[t] <- 1 - prod(1 - p[t,])
  }
# Likelihood
   for (i in 1:n.ind){
     z[i,first[i]] <- ch[i,first[i]]
# phi adjusted for probability of losing tag
   for (t in (first[i]+1):n.years){
     ##mu1[i,t] <- z[i,t-1] * phi[t-1]##
     mu1[i,t] <- z[i,t-1] * (phi[t-1] * XXX) #adjust this value (XXX) for prob of (not) losing tag(s)
     mu2[i,t] <- z[i,t] * gamma[t-1] * pstar[t]
     z[i,t] ~ dbern(mu1[i,t])
     ch[i,t] ~ dbern(mu2[i,t])
   }
  } # end likelihood
## pop abundances
# adjust this value (XXX) for prob of (not) losing tag(s)
   for(t in 1:n.years){
      Nin[t] <- (sum(ch[,t])/pstar[t]) * XXX
   }
} # end model
"
```

LITERATURE CITED

Aboim, M.A., Mavarez J., Bernatchez, L., & Coelho, M.M. (2010). Introgressive hybridization between two Iberian endemic cyprinid fish: a comparison between two independent hybrid zones. Journal of Evolutionary Biology, 23, 817-828.

Albert, J. S., Craig, J. M., Tagliacollo, V. A., & Petry, P. (2018). Upland and Lowland Fishes: A Test of the River Capture Hypothesis. In C. Hoorn, A. Perrigo, & A. Antonelli (Eds.), Mountains, Climate and Biodiversity (pp. 273–294).

Allen, D. C., & Vaughn, C. C. (2009). Burrowing behavior of freshwater mussels in experimentally manipulated communities. Journal of the North American Benthological Society, 28(1), 93–100. https://doi.org/10.1899/07-170.1

Allen, D. C., & Vaughn, C. C. (2010). Complex hydraulic and substrate variables limit freshwater mussel species richness and abundance. 29(February), 383–394. https://doi.org/10.1899/09-024.1

Allendorf, F.W., Leary, R.F., Spruell, P., & Wenburg, J.K. (2001). The problems with hybrids: setting conservation guidelines. Trends in Ecology and Evolution, 16, 613-622.

Alter, S. E., Munshi-South, J., & Stiassny, M. L. J. (2017). Genomewide SNP data reveal cryptic phylogeographic structure and microallopatric divergence in a rapids-adapted clade of cichlids from the Congo River. Molecular Ecology, 26(5), 1401–1419. https://doi.org/10.1111/mec.13973

Amyot, J.-P., & Downing, J. A. (1997). Seasonal variation in vertical and horizontal movement of the freshwater bivalve. Freshwater Biology, 37, 345–354.

April, J., Hanner, R. H., Mayden, R. L., & Bernatchez, L. (2013). Metabolic Rate and Climatic Fluctuations Shape Continental Wide Pattern of Genetic Divergence and Biodiversity in Fishes. PLoS ONE, 8(7). https://doi.org/10.1371/journal.pone.0070296

Arnold, M.L. & Martin, N.H. (2009). Adaptation by introgression. Journal of Biology 8, 82.

Arnold, M.L., Sapir, Y., & Martin, N.H. (2008). Genetic exchange and the origin of adaptations: prokaryotes to primates. Philosophical Transactions of the Royal Society B, 363, 2813-2820.

Ashton, M. J., Tiemann, J. S., & Hua, D. (2017). Evaluation of Costs Associated with Externally Affixing Pit Tags to Freshwater Mussels Using Three Commonly Employed Adhesives. Freshwater Mollusk Biology and Conservation, 20(2), 114. https://doi.org/10.31931/fmbc.v20i2.2017.114-122

Baker, V. R., & Penteado-Orellana, M. M. (1977). Adjustment to Quaternary Climatic Change by the Colorado River in Central Texas. The Journal of Geology, 85(4), 395–422. https://doi.org/10.1086/628315

Barlow, G.W. (1958). High salinity mortality of desert pupfish, Cyprinodon macularius. Copeia, 1958, 231-232.

Bartáková, V., Bryja, J., Šanda, R., Bektas, Y., Stefanov, T., Choleva, L., … Reichard, M. (2019). High cryptic diversity of bitterling fish in the southern West Palearctic. Molecular Phylogenetics and Evolution, 133, 1–11. https://doi.org/10.1016/j.ympev.2018.12.025

Barton, H.A. (1986). The effect of linkage and density-dependent regulation on gene flow. Heredity, 57, 415-426.

Beaumont, M. A. (2008). Joint determination of topology, divergence time, and immigration in population trees. In Simulations, Genetics and Human Prehistory.

Bennett, K. L., Kaddumukasa, M., Shija, F., Djouaka, R., Misinzo, G., Lutwama, J., … Walton, C. (2018). Comparative phylogeography of Aedes mosquitoes and the role of past climatic change for evolution within Africa. Ecology and Evolution, (June), 1–18. https://doi.org/10.1002/ece3.3668

Berendzen, P. B., Simons, A. M., & Wood, R. M. (2003). Phylogeography of the northern hogsucker, Hypentelium nigricans (Teleostei: Cypriniformes): genetic evidence for the existence of the ancient Teays River. Journal of Biogeography, 30(8), 1139–1152. https://doi.org/10.1046/j.1365-2699.2003.00888.x

Boavida, J. R., Becheler, R., Choquet, M., Frank, N., Taviani, M., Bourillet, J.-F., … Arnaud-Haond, S. (2019). Out of the Mediterranean? Post-glacial colonisation pathways varied among cold-water coral species. Journal of Biogeography, (March), 915–931. https://doi.org/10.1111/jbi.13570

Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. Bioinformatics, 26(10), 1372–1373. https://doi.org/10.1093/bioinformatics/btq110

Brussard, P. F. (1984). Geographic Patterns and Environmental Gradients: The Central-Marginal Model in Drosophila Revisited. Annual Review of Ecology and Systematics, 15(128), 25–64.

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & Roychoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. Molecular Biology and Evolution, 29(8), 1917–1932. https://doi.org/10.1093/molbev/mss086

Buerkle, C.A. & Lexer, C. (2008). Admixture as the basis for genetic mapping. Trends in Ecology and Evolution, 23, 686-694.

Burridge, C. P., Craw, D., & Waters, J. M. (2006). River capture, range expansion, and cladogensis: the genetic signature of freshwater vicariance. Evolution, 60(5), 1038. https://doi.org/10.1554/05-439.1

Burridge, Christopher P., Craw, D., Jack, D. C., King, T. M., & Waters, J. M. (2008). Does fish ecology predict dispersal across a river drainage divide? Evolution, 62(6), 1484–1499. https://doi.org/10.1111/j.1558-5646.2008.00377.x

Carvajal-Quintero, J. D., Escobar, F., Alvarado, F., Villa-Navarro, F. A., Jaramillo-Villa, Ú., & Maldonado-Ocampo, J. A. (2015). Variation in freshwater fish assemblages along a regional elevation gradient in the northern Andes, Colombia. Ecology and Evolution, 5(13), 2608–2620. https://doi.org/10.1002/ece3.1539

Carvajal-Quintero, J., Villalobos, F., Oberdorff, T., Grenouillet, G., Brosse, S., Hugueny, B., … Tedesco, P. A. (2019). Drainage network position and historical connectivity explain global patterns in freshwater fishes' range size. Proceedings of the National Academy of Sciences, 116(27), 13434–13439. https://doi.org/10.1073/pnas.1902484116

Cobb, D. G., Galloway, T. D., & Flannagan, J. F. (1992). Effects of discharge and substrate stability on density and species composition of stream insects. Canadian Journal of Fisheries and Aquatic Sciences, 49(9), 1788–1795. https://doi.org/10.1139/f92-198

Conner, J. V., & Suttkus, R. D. (1986). Zoogeography of Freshwater Fishes of the Western Gulf Slope of North America. In C. H. Hocutt & E. O. Wiley (Eds.), Zoogeography of North American Freshwater Fishes (pp. 413–456).

Cornuet, J. M., Santos, F., Beaumont, M. A., Robert, C. P., Marin, J. M., Balding, D. J., … Estoup, A. (2008). Inferring population history with DIY ABC: A user-friendly approach to approximate Bayesian computation. Bioinformatics, 24(23), 2713–2719. https://doi.org/10.1093/bioinformatics/btn514

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., … Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics, 30(8), 1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Coyne, J.A. & Orr, H.A. (2004). Speciation. Sunderland, MA, Sinauer Associates.

Craig, C. A., Kollaus, K. A., Behen, K. P. K., & Bonner, T. H. (2016). Relationships among spring flow, habitats, and fishes within evolutionary refugia of the Edwards Plateau. Ecosphere, 7(2). https://doi.org/10.1002/ecs2.1205

Crandall, R. M., Hayes, C. R., & Ackland, E. N. (2003). Application of the intermediate disturbance hypothesis to flooding. Community Ecology, 4(2), 225–232.

Datry, T., Melo, A. S., Moya, N., Zubieta, J., De la Barra, E., & Oberdorff, T. (2016). Metacommunity patterns across three Neotropical catchments with varying environmental harshness. Freshwater Biology, 61(3), 277–292. https://doi.org/10.1111/fwb.12702

Dias, M. S., Oberdorff, T., Hugueny, B., Leprieur, F., Jézéquel, C., Cornu, J. F., … Tedesco, P. A. (2014). Global imprint of historical connectivity on freshwater fish biodiversity. Ecology Letters, 17(9), 1130–1140. https://doi.org/10.1111/ele.12319

Driscoe, A. L., Nice, C. C., Busbee, R. W., Hood, G. R., Egan, S. P., & Ott, J. R. (2019). Host plant associations and geography interact to shape diversification in a specialist insect herbivore. Molecular Ecology, 28(18), 4197–4211. https://doi.org/10.1111/mec.15220

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7(1), 214. https://doi.org/10.1186/1471-2148-7-214

Dufresnes, C., Nicieza, A. G., Litvinchuk, S. N., Rodrigues, N., Jeffries, D. L., Vences, M., … Martínez-Solano, Í. (2020). Are glacial refugia hotspots of speciation and cyto-nuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. Molecular Ecology, (November 2019), 1–15. https://doi.org/10.1111/mec.15368

Dufresnes, C., Strachinis, I., Suriadna, N., Mykytynets, G., Cogălniceanu, D., Székely, P., … Denoël, M. (2019). Phylogeography of a cryptic speciation continuum in Eurasian spadefoot toads (Pelobates). Molecular Ecology, (February), 1–14. https://doi.org/10.1111/mec.15133

Duvernell, D. D., Westhafer, E., & Schaefer, J. F. (2019). Late Pleistocene range expansion of North American topminnows accompanied by admixture and introgression. Journal of Biogeography, (December 2018), jbi.13597. https://doi.org/10.1111/jbi.13597

Echelle, A. A., Carson, E. W., Echelle, A. F., Van Den Bussche, R. A., Dowling, T. E., & Meyer, A. (2005). Historical biogeography of the New-World pupfish genus Cyprinodon (Teleostei: Cyprinodontidae). Copeia, 2005(2), 320–339. https://doi.org/10.1643/CG-03-093R3

Echelle, A.A., Land, N.J., Borden, W.C., Schwemm, M.R., Hoagstrom, C.W., Eisenhour, D.J., Mayden, R.L., & Van Den Bussche, R.A. (2018). Molecular systematics of the North American chub genus Macrhybopsis (Teleostei: Cyprinidae). Zootaxa, 4375, 537-554.

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: The central-marginal hypothesis and beyond. Molecular Ecology, 17(5), 1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Molecular Ecology, 16, 463–475.

Eisenhour, D. J. (2004). Systematics, variation, and speciation of the Macrhybopsis aestivalis complex west of the Mississippi River. Bulletin of the Alabama Museum of Natural History, 23(January 2004), 9–48. https://doi.org/10.2307/1447972

Fagundes, N. J. R., Ray, N., Beaumont, M., Neuenschwander, S., Salzano, F. M., Bonatto, S. L., & Excoffier, L. (2007). Statistical evaluation of alternative models of human evolution. Proceedings of the National Academy of Sciences, 104(45), 17614–17619. https://doi.org/10.1073/pnas.0708280104

Falush, D., Stephens, M., & Pritchard, J.K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics, 164, 1567-1587.

Flecker, A. S., & Feifarek, B. (1994). Disturbance and the temporal variability of invertebrate assemblages in two Andean streams. Freshwater Biology, 31(2), 131–142. https://doi.org/10.1111/j.1365-2427.1994.tb00847.x

Fletcher, N. K., Acevedo, P., Herman, J. S., Paupério, J., Alves, P. C., & Searle, J. B. (2019). Glacial cycles drive rapid divergence of cryptic field vole species. Ecology and Evolution, (October 2019), 1–13. https://doi.org/10.1002/ece3.5846

Franssen, N. R., Gido, K. B., Guy, C. S., Tripe, J. A., Shrank, S. J., Strakosh, T. R., … Paukert, C. P. (2006). Effects of floods on fish assemblages in an intermittent prairie stream. Freshwater Biology, 51(11), 2072–2086. https://doi.org/10.1111/j.1365-2427.2006.01640.x

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics, 28, 3150-3152.

Galloway, W. E., Whiteaker, T. L., & Ganey-Curry, P. (2011). History of Cenozoic North American drainage basin evolution, sediment yield, and accumulation in the Gulf of Mexico basin. Geosphere, 7(4), 938–973. https://doi.org/10.1130/GES00647.1

Gangloff, M. M., & Feminella, J. W. (2007). Stream channel geomorphology influences mussel abundance in southern Appalachian streams, U.S.A. Freshwater Biology, 52(1), 64–74. https://doi.org/10.1111/j.1365-2427.2006.01673.x

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. Statistical Science, 7(4), 457–511. https://doi.org/10.2307/2246134

Gilbert, C. R., Mayden, R. L., & Powers, S. L. (2017). Morphological and genetic evolution in eastern populations of the Macrhybopsis aestivalis complex (Cypriniformes: Cyprinidae), with the descriptions of four new species. 4247(5), 501–555.

Gilbert, K.J., Andrew, R.L., Bock, D.G., Franklin, M.T., Kane, N.C., Moore, J.S., Moyers, B.T., Renaut, S., Rennison, D.J., Veen, T. & Vines, T.H. (2012). Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. Molecular Ecology, 21(20), 4925-4930.

Gompert, Z., & Buerkle, C.A. (2009). A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. Molecular Ecology, 18: 1207-1224.

Gompert, Z., & Buerkle, C.A. (2011). Bayesian estimation of genomic clines. Molecular Ecology, 20, 2111-2127.

Gompert, Z., & Buerkle, C.A. (2012). Bgc: software for Bayesian estimation of genomic clines. Molecular Ecology Resources, 12, 1168-1176.

Gompert, Z., & Buerkle, C.A. (2016). What, if anything, are hybrids: enduring truths and larkiaes associated with population structure and gene flow. Evolutionary Applications, 9, 909-923.

Gompert, Z., Lucas, L. K., Buerkle, C. A., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2014). Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. Molecular Ecology, 23(18), 4555–4573. https://doi.org/10.1111/mec.12811

Gompert, Z., Lucas, L.K., Nice, C.C., Fordyce, J.A., Forister, M.L., & Buerkle, C.A. (2012). Genomic regions with a history of divergent selection affect fitness of hybrids between two butterfly species. Evolution, 66, 2167-2181.

Goodding, D. D., Williams, M. G., Ford, D. F., Williams, L. R., & Ford, N. B. (2019). Associations between substrate and hydraulic variables and the distributions of a sculptured and an unsculptured unionid mussel. Freshwater Science, 38(3), 543–553. https://doi.org/10.1086/704795

Grant, V. (1981). Plant speciation. New York, USA: Columbia University Press.

Grimm, N. B., & Fisher, S. G. (1989). Stability of Periphyton and Macroinvertebrates to Disturbance by Flash Floods in a Desert Stream. Journal of the North American Benthological Society, 8(4), 293–307. https://doi.org/10.2307/1467493

Groemping, U. (2006) Relative Importance for Linear Regression in R: The Package relaimpo. Journal of Statistical Software, 17, 1-27.

Haq, B. U., Hardenbol, J., & Vail, P. R. (1987). Chronology of Fluctuating Sea Levels Since the Triassic. Science, 235(4793), 1156–1167. https://doi.org/10.1126/science.235.4793.1156

Harrison, R.G. (1990). Hybrid zones: windows on evolutionary process. Oxford Surveys in Evolutionary Biology, 7, 69-128.

Haselhorst, M. S. H., Parchman, T. L., & Buerkle, C. A. (2019). Genetic evidence for species cohesion, substructure, and hybrids in spruce. Molecular Ecology, (June 2017), 1–17. https://doi.org/10.1111/mec.15056

Hasselman, D. J., Argo, E. E., McBride, M. C., Bentzen, P., Schultz, T. F., Perez-Umphrey, A.A., & Palkovacs, E. P. (2014). Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. Molecular Ecology, 23, 1137–1152.

Hastie, L. C., Boon, P. J., Young, M. R., & Way, S. (2001). The effects of a major flood on an endangered freshwater mussel population. Biological Conservation, 98(1), 107–115. https://doi.org/10.1016/S0006-3207(00)00152-X

Healey, A. J. E., McKeown, N. J., Taylor, A. L., Provan, J., Sauer, W., Gouws, G., & Shaw, P. W. (2018). Cryptic species and parallel genetic structuring in Lethrinid fish: Implications for conservation and management in the southwest Indian Ocean. Ecology and Evolution, (December 2017), 2182–2195. https://doi.org/10.1002/ece3.3775

Heath, D., Bettles, C.M., & Roff, D. (2010). Environmental factors associated with reproductive barrier breakdown in sympatric trout populations on Vancouver Island. Evolutionary Applications, 3, 77-90.

Higgins, C.L., & Wilde, G.R. (2005). The role of salinity in structuring fish assemblages in a prairie stream system. Hydrobiologia, 549, 197-203.

Hornbach, D. J., Kurth, V. J., & Hove, M. C. (2010). Variation in Freshwater Mussel Shell Sculpture and Shape Along a River Gradient. The American Midland Naturalist, 164(1), 22–36. https://doi.org/10.1674/0003-0031-164.1.22

Hubbs, C.L. (1955). Hybridization between fish species in nature. Systematic Zoology, 4, 1-20.

Hughes, L. C., Cardoso, Y. P., Sommer, J. A., Cifuentes, R., Cuello, M., Somoza, G. M., … Ortí, G. (2020). Biogeography, habitat transitions and hybridization in a radiation of South American silverside fishes revealed by mitochondrial and genomic RAD data. Molecular Ecology, (December 2019), 1–14. https://doi.org/10.1111/mec.15350

Inoue, K., Harris, J. L., Robertson, C. R., Johnson, N. A., & Randklev, C. R. (2019). A comprehensive approach uncovers hidden diversity in freshwater mussels (Bivalvia: Unionidae) with the description of a novel species. Cladistics, 36, 88–113. https://doi.org/10.1111/cla.12386

Inoue, K., Levine, T. D., Lang, B. K., & Berg, D. J. (2014). Long-term mark-and-recapture study of a freshwater mussel reveals patterns of habitat use and an association between survival and river discharge. Freshwater Biology, 59(9), 1872–1883. https://doi.org/10.1111/fwb.12389

Janes, J.K., Miller, J.M., Dupuis, J.R., Malenfant, R.M., Gorrell, J.C., Cullingham, C.I. & Andrew, R.L. (2017). The K= 2 conundrum. Molecular Ecology, 26(14), 3594-3602.

Jelks, H. L., Walsh, S. J., Burkhead, N. M., Contreras-Balderas, S., Díaz-Pardo, E., Hendrickson, D.A., Lyons, J., Mandrak, N. E., McCormick, F., Nelson, J. S., Platania, S. P., Porter, B. A., Renaud, C. B., Jacobo Schmitter-Soto, J., Taylor, E. B., & Warren, Jr, M.L. (2008). Conservation status of imperiled North American freshwater and diadromous fishes. Fisheries, 33, 372-407.

Jenkins, C. N., Van Houtan, K. S., Pimm, S. L., & Sexton, J. O. (2015). US protected lands mismatch biodiversity priorities. Proceedings of the National Academy of Sciences, 112(16), 5081–5086. https://doi.org/10.1073/pnas.1418034112

Jetz, W., Rahbek, C., & Colwell, R. K. (2004). The coincidence of rarity and richness and the potential signature of history in centres of endemism. Ecology Letters, 7(12), 1180–1191. https://doi.org/10.1111/j.1461-0248.2004.00678.x

Johnson, N. A., Smith, C. H., Pfeiffer, J. M., Randklev, C. R., Williams, J. D., & Austin, J. D. (2018). Integrative taxonomy resolves taxonomic uncertainty for freshwater mussels being considered for protection under the U.S. Endangered Species Act. Scientific Reports, 8(1), 1–16. https://doi.org/10.1038/s41598-018-33806-z

Jonkman, S. N., & Kelman, I. (2005). An analysis of the causes and circumstances of flood disaster deaths. Disasters, 29(1), 75–97.

Junk, W. J., Bayley, P. B., & Sparks, R. E. (1989). The flood pulse concept. International Large River Symposium, (September 1989), 110–127.

Kingston, S.E., Parchman, T.L., Gompert, Z., Buerkle, C.A., & Braun, M.J. (2017). Heterogeneity and concordance in locus-specific differentiation and introgression between species of towhees. Journal of Evolutionary Biology, 30, 474-485.

Konkle, B.R., & Philipp, D.P. (1992). Asymmetric hybridization between two species of sunfishes (Lepomis: Centrarchidae). Molecular Ecology, 1: 215-222.

Konrad, C. P. (2003). Effects of Urban Development on Floods. U.S. Geological Survey, d(November), 1–4. https://doi.org/USGS Fact Sheet FS-076-03

Kurth, J., Loftin, C., Zydlewski, J., & Rhymer, J. (2007). PIT tags increase effectiveness of freshwater mussel recaptures. Benthol. Soc, 26(2), 253–260. https://doi.org/10.1899/0887-3593(2007)26[253:PTIEOF]2.0.CO;2

Lamaze, F.C., Sauvage, C., Marie, A., Garant, D., & Bernatchez, L. (2012). Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (Salvelinus fontinalis). Molecular Ecology, 21, 2877-2895.

Lamp, T., & Avise, J.C. (1986). Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: the influence of mating behavior. Proceedings of the National Academy of Sciences, USA, 83: 2526-2530.

Langmead, B., & Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9, 357-359.

Le Corre, V., & Kremer, A. (1998). Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. Journal of Evolutionary Biology, 11(4), 495–512. https://doi.org/10.1007/s000360050102

Lepais, O., Petit, R.J., Guichoux, E., Lavabre, J.E., Alberto, F., Kremer, A., & Gerber, S. (2009). Species relative abundance and direction of introgression in oaks. Molecular Ecology, 18, 2228-2242.

Leroy, B., Dias, M. S., Giraud, E., Hugueny, B., Jézéquel, C., Leprieur, F., … Tedesco, P. A. (2019). Global biogeographical regions of freshwater fish species. Journal of Biogeography, (July), 2407–2419. https://doi.org/10.1111/jbi.13674

Levin, D.A. (2002). The role of chromosomal change in plant evolution. Oxford: Oxford University Press.

Levin, D.A. (1983). Polyploidy and novelty in flowering plants. American Naturalist, 122, 1–25.

Levine, T. D., Hansen, H. B., & Gerald, G. W. (2014). Effects of shell shape, size, and sculpture in burrowing and anchoring abilities in the freshwater mussel Potamilus alatus (Unionidae). Biological Journal of the Linnean Society, 111(1), 136–144. https://doi.org/10.1111/bij.12178

Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14): 1754-1760.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27, 2987–2993.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22, 1658-1659.

Machado, C. B., Galetti, P. M., & Carnaval, A. C. (2018). Bayesian analyses detect a history of both vicariance and geodispersal in Neotropical freshwater fishes. Journal of Biogeography, 45(6), 1313–1325. https://doi.org/10.1111/jbi.13207

Maloney, K. O., Lellis, W. A., Bennett, R. M., & Waddle, T. J. (2012). Habitat persistence for sedentary organisms in managed rivers: The case for the federally endangered dwarf wedgemussel (Alasmidonta heterodon) in the Delaware River. Freshwater Biology, 57(6), 1315–1327. https://doi.org/10.1111/j.1365-2427.2012.02788.x

Maltchik, L., & Pedro, F. (2001). Responses of Aquatic Macrophytes to Disturbance by Flash Floods in a Brazilian Semiarid Intermittent Stream1. Biotropica, 33(4), 566. https://doi.org/10.1646/0006-3606(2001)033[0566:roamtd]2.0.co;2

Mandeville, E.G., Parchman, T.L., McDonald, D.B., & Buerkle, C.A. (2015). Highly variable reproductive isolation among pairs of Catostomus species. Molecular Ecology, 24, 1856-1872.

Marie, A.D., Bernatchez, L., & Garant, D. (2012). Environmental factors correlate with hybridization in stocked brook charr (Salvelinus fontinalis). Canadian Journal of Fisheries and Aquatic Sciences, 69, 884-893.

Martin, N.H., Bouck, A.C. & Arnold, M.L. (2006). Detecting adaptive trait introgression between Iris fulva and I. brevicaulis in highly selective field conditions. Genetics, 172, 2481-2489.

Martin, N.H., Bouck, A.C., & Arnold, M.L. (2005). Loci affecting long-term hybrid survivorship in Louisiana irises: implications for reproductive isolation and introgression. Evolution, 59, 2116-2124.

Martinson, D. G., Pisias, N. G., Hays, J. D., Imbrie, J., Moore, T. C., & Shackleton, N. J. (1987). Age Dating and the Orbital Theory of the Ice Ages: Development of a High-Resolution 0 to 300,000-Year Chronostratigraphy. Quaternary Research, 27(1), 1–29. https://doi.org/10.1016/0033-5894(87)90046-9

Mayden, R. L. (1985). Biogeography of Ouachita Highland Fishes. The Southwestern Naturalist, 30(2), 195–211.

Mayden, R. L. (1988). Vicariance Biogeography, Parsimony, and Evolution in North American Freshwater Fishes. 37(4), 329–355.

McGaughran, A., Terauds, A., Convey, P., & Fraser, C. I. (2019). Genome-wide SNP data reveal improved evidence for Antarctic glacial refugia and dispersal of terrestrial invertebrates. Molecular Ecology, 81(11), 1166–1173. https://doi.org/10.1111/jofi.12047.

McMullen, L. E., & Lytle, D. A. (2012). Quantifying invertebrate resistance to floods: A global-scale meta-analysis. Ecological Applications, 22(8), 2164–2175. https://doi.org/10.1890/11-1650.1

Meador, J. R., Peterson, J. T., & Wisniewski, J. M. (2011). An evaluation of the factors influencing freshwater mussel capture probability, survival, and temporary emigration in a large lowland river. Journal of the North American Benthological Society, 30(2), 507–521. https://doi.org/10.1899/10-105.1

Meirmans, P.G. (2015). Seven common mistakes in population genetics and how to avoid them. Molecular Ecology, 24(13), 3223-3231.

Meirmans, P.G. (2019). Subsampling reveals that unbalanced sampling affects STRUCTURE results in a multi-species dataset. Heredity, 122, 276-287.

Melo, B. F., Ochoa, L. E., Vari, R. P., & Oliveira, C. (2016). Cryptic species in the Neotropical fish genus Curimatopsis (Teleostei, Characiformes). Zoologica Scripta, 45(6), 650–658. https://doi.org/10.1111/zsc.12178

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols, 2010.6: pdb-prot5448.

Morales, Y., Weber, L. J., Mynett, A. E., & Newton, T. J. (2006). Effects of substrate and hydrodynamic conditions on the formation of mussel beds in a large river. Journal of the North American Benthological Society, 25(3), 664–676. https://doi.org/10.1899/0887-3593(2006)25[664:eosahc]2.0.co;2

Nadeau, N. (2014). Butterfly genomics sheds light on the process of hybrid speciation. Molecular Ecology, 23, 4441-4443.

Nagle, B. C., & Simons, A. M. (2012). Rapid diversification in the North American minnow genus Nocomis. Molecular Phylogenetics and Evolution, 63(3), 639–649. https://doi.org/10.1016/j.ympev.2012.02.013

NatureServe. 2010. Digital Distribution Maps of the Freshwater Fishes in the Conterminous United States. Version 3.0. Arlington, VA. U.S.A.

Nei, M. (1987). Molecular evolutionary genetics. Columbia University Press, New York.

Newton, T. J., Zigler, S. J., & Gray, B. R. (2015). Mortality, movement and behaviour of native mussels during a planned water-level drawdown in the Upper Mississippi River. Freshwater Biology, 60(1), 1–15. https://doi.org/10.1111/fwb.12461

Nichols, J. D., & Pollock, K. H. (1990). Estimation of recruitment from immigration versus in situ reproduction using Pollock's robust design. Ecology, 71(1), 21–26. https://doi.org/10.2307/1940243

Nolte, A.W., Gompert, Z., & Buerkle, C.A. (2009). Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. Molecular Ecology, 18, 2615-2627.

Oberdorff, T., Lek, S., & Guegan, J.-F. (1999). Patterns of endemism in riverine fish of the Northern Hemisphere. Ecology Letters, 2(2), 75–81. https://doi.org/10.1046/j.1461-0248.1999.t01-2-22051.x

Oksanen, J.F., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., & Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-5.

Osborne, M. J., Diver, T. A., Hoagstrom, C. W., & Turner, T. F. (2016). Biogeography of "Cyprinella lutrensis": Intensive genetic sampling from the Pecos River "melting pot" reveals a dynamic history and phylogenetic complexity. Biological Journal of the Linnean Society, 117(2), 264–284. https://doi.org/10.1111/bij.12664

Ostberg, C.O., Slatton, S.L., & Rodriguez, R.J. (2004). Spatial partitioning and asymmetric hybridization among sympatric coastal steelhead trout (Oncorhynchus mykiss irideus), coastal cutthroat trout (O. clarki clarki) and interspecific hybrids. Molecular Ecology, 13, 2773-2788.

Ostrand, K.G. (2000). Abiotic determinants of fish assemblage structure in the upper Brazos River, Texas. Doctoral dissertation. Texas Tech University, Lubbock, Texas.

Ostrand, K.G., & Wilde, G.R. (2001). Temperature, dissolved oxygen, and salinity tolerances of five prairie stream fishes and their role in explaining fish assemblage patterns. Transactions of the American Fisheries Society, 130, 742-749.

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. Molecular Ecology, 21(12), 2991–3005. https://doi.org/10.1111/j.1365-294X.2012.05513.x

Parchman, T.L., Gompert, Z., Braun, M.J., Brumfield, R.T., McDonald, D.B., Uy, J.A.C., Zhang, G., Jarvis, E.D., Schlinger, B.A., & Buerkle, C.A. (2013). The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. Molecular Ecology, 22, 3304–3317.

Payseur, B.A. (2010). Using differential introgression in hybrid zones to identify genomic regions involved in speciation. Molecular Ecology Resources, 10, 806-820.

Pérez-Rodríguez, R., Domínguez-Domínguez, O., Mar-Silva, A. F., Doadrio, I., & Pérez-Ponce de León, G. (2016). The historical biogeography of the southern group of the sucker genus Moxostoma (Teleostei: Catostomidae) and the colonization of central Mexico. Zoological Journal of the Linnean Society, 177(3), 633–647. https://doi.org/10.1111/zoj.12383

Perkin, J.S., Gido, K.B., Costigan, K.H., Daniels, M.D., & Johnson, E.R. (2015). Fragmentation and drying ratchet down Great Plains stream fish diversity. Aquatic Conservation: Marine and Freswhater Ecosystems, 25, 639-655.

Peters, D. L., Caissie, D., Monk, W. A., Rood, S. B., & St-Hilaire, A. (2016). An ecological perspective on floods in Canada. Canadian Water Resources Journal, 41(1–2), 288–306. https://doi.org/10.1080/07011784.2015.1070694

Poff, N. L. R., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., … Stromberg, J. C. (1997). The natural flow regime: A paradigm for river conservation and restoration. BioScience, 47(11), 769–784. https://doi.org/10.2307/1313099

Pollock, K. H. (1982). A Capture-Recapture Design Robust to Unequal Probability of Capture. The Journal of Wildlife Management, 46(3), 752–757.

Power, M. E., & Stewart, A. J. (1987). Disturbance and Recovery of an Algal Assemblage Following Flooding in an Oklahoma Stream. The American Midland Naturalist, 117(2), 333–345.

Pritchard, J.K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155, 945-959.

Puritz, J. B., Hollenbeck, C.M., & Gold, J. R. (2014). dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431; DOI 10.7717/peerj.431.

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Randklev, C. R., Hart, M. A., Khan, J. M., Tsakiris, E. T., & Robertson, C. R. (2019). Hydraulic requirements of freshwater mussels (Unionidae) and a conceptual framework for how they respond to high flows. Ecosphere, 10(2). https://doi.org/10.1002/ecs2.2975

Reinert, T. R., Wallin, J., Griffin, M. C., Conroy, M. J., & Avyle, M. J. Van Den. (1998). Long-term retention and detection of oxytetracycline marks applied to hatchery-reared larval striped bass, Morone saxatilis. Canadian Journal of Fisheries and Aquatic Sciences, 55(1998), 539–543.

Rhymer, J.M., & Simberloff, D. (1996). Extinction by hybridization and introgression. Annu. Rev. Ecol. Syst, 27, 83-109.

Riecke, T. V., Leach, A. G., Gibson, D., & Sedinger, J. S. (2018). Parameterizing the robust design in the BUGS language: Lifetime carry-over effects of environmental conditions during growth on a long-lived bird. Methods in Ecology and Evolution, 2018(April), 1–12. https://doi.org/10.1111/2041-210X.13065

Rieseberg, L. H., Raymond, O., Rosenthal, D. M., Lai, Z., Livingstone, K., Nakazato, T., … & Lexer, C. (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. Science, 301(5637), 1211-1216.

Riggs, S. R. (1984). Paleoceanographic Model of Neogene Phosphorite Deposition, U.S. Atlantic Continental Margin. Science, 223(4632), 123–131. https://doi.org/10.1126/science.223.4632.123

Robinson, C. T. (2012). Long-term changes in community assembly, resistance, and resilience following experimental floods. Ecological Applications, 22(7), 1949–1961. https://doi.org/10.1890/11-1042.1

Ruppel, D.S., Sotola, V.A., Gurbux, O.A., Martin, N.H., & Bonner, T.H. (2017). RFP No. 212f for Endangered species Research Projects for the prairie chub. Final Report.

Sansom, B. J., Bennett, S. J., Atkinson, J. F., & Vaughn, C. C. (2018). Long - term persistence of freshwater mussel beds in labile river channels. Freshwater Biology, 63(June), 1469–1481. https://doi.org/10.1111/fwb.13175

Santucci Jr., V.J., Gephard, S.R., & Pescitelli, S.M. (2005). Effects of multiple low-head dams on fish, macroinvertebrates, habitat, and water quality in the Fox River, Illinois. North American Journal of Fisheries Management, 25, 975-992.

Schaefer, J., Duvernell, D., & Campbell, D.C. (2016). Hybridization and introgression in two ecologically dissimilar Fundulus hybrid zones. Evolution, 70, 1051-1063.

Schmera, D., Árva, D., Boda, P., Bódis, E., Bolgovics, Á., Borics, G., … Erős, T. (2018). Does isolation influence the relative role of environmental and dispersal-related processes in stream networks? An empirical test of the network position hypothesis using multiple taxa. Freshwater Biology, 63(1), 74–85. https://doi.org/10.1111/fwb.12973

Sepkoski, J. J., & Rex, M. A. (1974). Distribution of Freshwater Mussels: Coastal Rivers as Biogeographic Islands. Society of Systematic Biologists, 23(2), 165–188.

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. Annual review of plant biology, 60, 561-588.

Sotola, V. A., Ruppel, D. S., Bonner, T. H., Nice, C. C., & Martin, N. H. (2019). Asymmetric introgression between fishes in the Red River basin of Texas is associated with variation in water quality. Ecology and Evolution, 9(4), 2083–2095. https://doi.org/10.1002/ece3.4901

Stobie, C. S., Oosthuizen, C. J., Cunningham, M. J., & Bloomer, P. (2018). Exploring the phylogeography of a hexaploid freshwater fish by RAD sequencing. Ecology and Evolution, (May 2017), 2326–2342. https://doi.org/10.1002/ece3.3821

Strayer, D. L. (1999). Use of Flow Refuges by Unionid Mussels in Rivers Author (s): David L . Strayer Source : Journal of the North American Benthological Society , Vol . 18 , No . 4 (Dec ., 1999), Published by : The University of Chicago Press on behalf of the Society for. The North American Benthological Society, 18(4), 468–476.

Strayer, D. L., & Smith, D. R. (2003). A Guide to Sampling Freshwater Mussel Populations. In American Fisheries Society. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=keh&AN=29339654&site=ehost-live

Sung, C.J., Bell, K.L., Nice, C.C., & Martin, N.H. (2018). Integrating Bayesian genomic cline analyses and association mapping of morphological and ecological traits to dissect reproductive isolation and introgression in a Louisiana Iris hybrid zone. Molecular Ecology, 27: 959-978.

Sylvia, D. A., & Galloway, W. E. (2006). Morphology and stratigraphy of the late Quaternary lower Brazos valley: Implications for paleo-climate, discharge and sediment delivery. Sedimentary Geology, 190(1–4), 159–175. https://doi.org/10.1016/j.sedgeo.2006.05.023

Talbot, C. J., Bennett, E. M., Cassell, K., Hanes, D. M., Minor, E. C., Paerl, H., … Xenopoulos, M. A. (2018). The impact of flooding on aquatic ecosystem services. Biogeochemistry, 141(3), 439–461. https://doi.org/10.1007/s10533-018-0449-7

Tedesco, P. A., Leprieur, F., Hugueny, B., Brosse, S., Dürr, H. H., Beauchard, O., … Oberdorff, T. (2012). Patterns and processes of global riverine fish endemism. Global Ecology and Biogeography, 21(10), 977–987. https://doi.org/10.1111/j.1466-8238.2011.00749.x

Teeter, K.C., L.M. Thibodeau, Z. Gompert, C.A. Buerkle, M.W. Nachman, & Tucker, P.K. (2010). The variable genomic architecture of isolation between hybridizing species of house mouse. Evolution, 64, 472–485.

Texas Parks and Wildlife Department. 2012. Texas Conservation Action Plan 2012-2016: Overview. Editor, Wendy Connally, Texas Conservation Action Plan Coordinator. Austin, TX.

U.S. Army Corps of Engineers. (2012). Red River Chloride Control Project, Texas and Okahoma. Fact Sheet as of February 6, 2012.

Underwood, D. M., Echelle, A. A. F., Eisenhour, D. J., Jones, M. D., Echelle, A. A. F., & Fisher, W. L. (2003). Genetic variation in western members of the Macrhybopsis aestivalis complex (Teleostei: Cyprinidae), with emphasis on those of the Red and Arkansas River basins. Copeia, 2003(3), 493–501. https://doi.org/10.1643/CG-02-033R1

Walker, M. J., Stockman, A. K., Marek, P. E., & Bond, J. E. (2009). Pleistocene glacial refugia across the appalachian mountains and coastal plain in the millipede genus Narceus: Evidence from population genetic, phylogeographic, and paleoclimatic data. BMC Evolutionary Biology, 9(1), 1–11. https://doi.org/10.1186/1471-2148-9-25

Watters, G T. (1994). Form and function of unionoidean shell sculpture and shape (Bivalvia). American Malacological Bulletin, 11(1), 1–20.

Watters, G. Thomas, O'Dee, S. H., & Chordas, S. (2001). Patterns of vertical migration in freshwater mussels (Bivalvia: Unionoida). Journal of Freshwater Ecology, 16(4), 541–549. https://doi.org/10.1080/02705060.2001.9663845

Wentworth, C. K. (1922). A Scale of Grade and Class Terms for Clastic Sediments. The Journal of Geology, 30(5), 377–392. https://doi.org/10.1086/622910

Whitney, K.D., Randell, R.A., Rieseber, L.H. (2006). Adaptive introgression of herbivore resistance traits in the weedy sunflower Helianthus annuus. American Naturalist, 167, 794-807.

Wilde, G.R. (2016). Migration of Arkansas river shiner and other broadcast spawning fishes in the Canadian River, New Mexico-Texas. Great Plains Landscape Conservation Cooperative Final Report.

Wilkinson, B. H., & Basse, R. A. (1978). Late Holocene history of the Central Texas coast from Galveston Island to Pass Cavallo. Bulletin of the Geological Society of America, 89(10), 1592–1600.

Wisniewski, J. M., Shea, C. P., Abbott, S., & Stringfellow, R. C. (2013). Imperfect Recapture: A Potential Source of Bias in Freshwater Mussel Studies. The American Midland Naturalist, 170(2), 229–247. https://doi.org/10.1674/0003-0031-170.2.229

Wood, T.E., Takebayashi, N., Barker, M.S., Mayrose, I., Greenspoon, P.B., & Rieseberg, L. (2009). The frequency of polyploid speciation in vascular plants. Proceedings of the National Academy of Sciences, USA, 106, 13875–13979.

Worthington, T.A., Echelle, A.A., Perkin, J.S., Mollenhauer, R., Farless, N., Dyer, J.J., Logue, D., & Brewer, S.K. (2016). The emblematic minnows of North American Great Plains: a synthesis of threats and conservation opportunities. Fish and Fisheries, 19, 271-307.

Wu, C. (2001). The genic view of the process of speciation. Journal of Evolutionary Biology, 14, 851-865.

Yau, M.M., & Taylor, E.B. (2013). Environmental and anthropogenic correlates of hybridization between westslope cutthroat trout (Oncorhynchus larkia lewisi) and introduced rainbow trout (O. mykiss). Conservation Genetics, 14, 885-900.

Young, S. P., & Isely, J. J. (2008). Evaluation of methods for attaching PIT tags and biotelemetry devices to freshwater mussels. Molluscan Research, 28(3), 175–178.

Yuri, T., Jernigan, R.W., Brumfield, R.T., Bhagabati, N.K. & Braun, M.J. (2009). The effect of marker choice on estimated levels of introgression across an avian (Pipridae: Manacus) hybrid zone. Molecular Ecology, 18, 4888–4903.

Zalapa, J.E., Brunet, J., & Guries, R.P. (2010). The extent of hybridization and its impact on the genetic diversity and population structure of an invasive tree, Ulmus pumila (Ulmaceae). Evolutionary Applications, 3, 157-168.

Zigler, S. J., Newton, T. J., Steuer, J. J., Bartsch, M. R., & Sauer, J. S. (2008). Importance of physical and hydraulic characteristics to unionid mussels: A retrospective analysis in a reach of large river. Hydrobiologia, 598(1), 343–360. https://doi.org/10.1007/s10750-007-9167-1.