Received: 13 November 2018 | Revised: 5 December 2018 | Accepted: 13 December 2018

DOI: 10.1002/ece3.4901

ORIGINAL RESEARCH

WILEY Ecology and Evolution Open Access

Asymmetric introgression between fishes in the Red River basin of Texas is associated with variation in water quality

V. Alex Sotola | David S. Ruppel | Timothy H. Bonner | Chris C. Nice | Noland H. Martin

Biology Department, Texas State University, San Marcos, Texas

Correspondence
V. Alex Sotola, Biology Department, Texas State University, San Marcos, TX.
Email: vas66@txstate.edu

Funding information
Texas Comptroller of Public Accounts, Grant/Award Number: RFP No. 212f

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. We integrated population genomic data and environmental data to investigate species divergence and the maintenance of species boundaries in these two species. We 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.

KEYWORDS
fish, genetic structure, hybridization, introgression, water quality variation

1 | 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 & 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,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

or even the slow extirpation of taxa through genetic swamping (Allendorf, Leary, Spruell, & Wenburg, 2001; Edmands, 2007), hybridization can also act as an evolutionarily creative force leading to increased genetic diversity, adaptive introgression, and even hybrid speciation (Arnold & Martin, 2009; Arnold et al., 2008; Gompert & Buerkle, 2016; Martin, Bouck, & Arnold, 2005, 2006; Rieseberg et al., 2003; Soltis & Soltis, 2009). 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 nonmodel 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 et al., 2015; Sung, Bell, Nice, & Martin, 2018). Genome-wide studies on nonmodel 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 have often been attributed to range shifts, mostly due to anthropogenic changes in the systems where these individuals occur (Perkin, Gido, Costigan, Daniels, & Johnson, 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 (Heath, Bettles, & Roff, 2010; Lamaze, Sauvage, Marie, Garant, & Bernatchez, 2012; Marie et al., 2012). Additionally, changes in environmental variables (e.g. water quality), which can be altered via natural causes or anthropogenic disturbances, can increase the amount of hybridization between species (Marie et al., 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 et al., 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). 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 (Echelle et al., 2018; Eisenhour, 2004). 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; Texas Parks & Wildlife Department, 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 (Echelle et al., 2018; Underwood et al., 2003). 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, we 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: (a) quantify patterns of genetic structure and the geographic and genomic extent of hybridization, (b) 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 (c) determine the degree to which water quality parameters are associated with genetic structuring.

2 | MATERIALS AND METHODS

2.1 | Sample collection

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; Table 1). Seines 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.

2.2 | DNA sequence generation, assembly, and variation

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 (Gompert et al., 2012; Mandeville et al., 2015; Meyer & Kircher, 2010; Parchman et al., 2012). 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) and sequenced over two lanes on an Illumina HiSeq 4000 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,

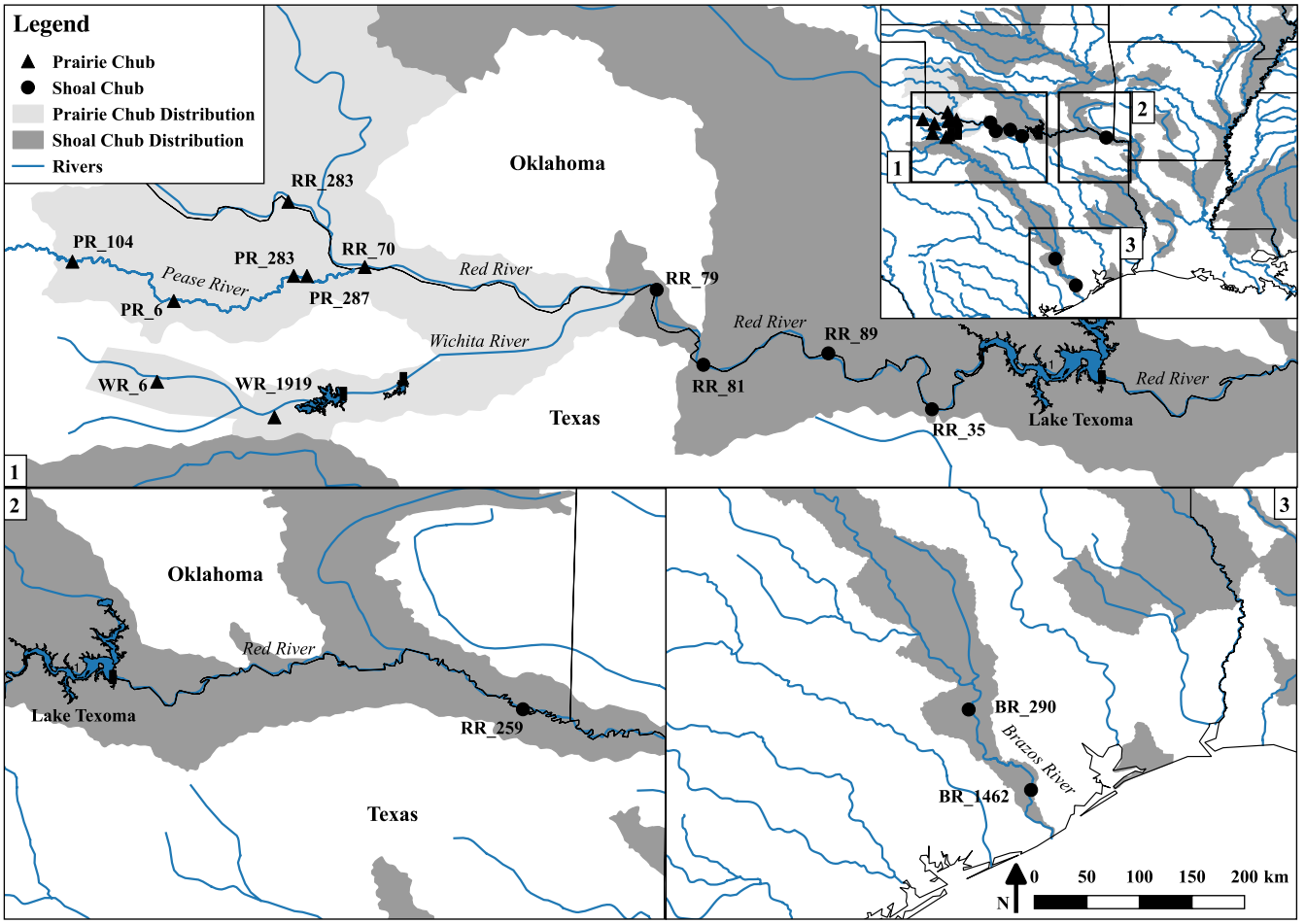

FIGURE 1 Map of locales where shoal chubs and prairie chubs were collected on the Red River (RR), Pease River (PR), Wichita River (WR), and Brazos River (BR). The map inset at the top right depicts the broad sampling frame statewide. Shapes denoting the sampling locations represent purportedly shoal chubs (circles) and prairie chubs (triangles) based on meristic morphological assignments. Light gray shading represents the prairie chub distribution, dark gray shading represents the shoal chub distribution (Data provided by NatureServe, 2010). Site codes are defined in Table 1

TABLE 1 Collection locales of shoal chubs and prairie chubs with sampling code, river system, road crossing, latitude and longitude, sample size (*N*), and nucleotide diversity (π)

Code	River	Road crossing	Latitude	Longitude	*N*	π
BR_1462	Brazos River	1,462	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

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; Li & Godzik, 2006) 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.

2.3 | Genetic structuring and diversity

To examine the genetic structuring of the shoal and prairie chubs, population genetic parameters were estimated using Entropy (Gompert et al., 2014a, 2014b; Mandeville et al., 2015). Entropy is a hierarchical model whereby an individual's assignment probability to each of any number of preassigned populations is estimated in a Bayesian framework. While interpretation of the output is similar to that of Structure (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000), 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).

2.4 | Genetic and environment associations

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 (°C), and dissolved oxygen (mg/L) were measured with a YSI 556 multiprobe 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 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 (percent R^2 explained) 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.

2.5 | Admixture class

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 are 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, we 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. We ran two independent MCMC analyses with 15,000 iterations, sampled every 5th iteration after a burn-in of 5,000 iterations. Admixture classes were estimated from samples of both independent MCMC analyses.

2.6 | Symmetry of introgression

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). Only individuals with Q_{12} (interpopulation 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

TABLE 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

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.

3 | RESULTS

3.1 | Field and genomic sampling

Sample sizes of fishes captured from the 15 sampling locales ranged from seven to 70 (Table 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.

3.2 | Genetic structuring and diversity

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). 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 2 and Appendix Table S1).

Principal component I explained 39.76% of the variation, and principal component II explained 13.51% of the variation in the genotypic data (Figure 2). Three primary clusters were indicated by PCA: (a) shoal chubs from the Brazos River, (b) prairie chubs from the Pease, Wichita, and upper Red River sites, and (c) shoal chubs from the downstream Red River sites. PC I segregated shoal chubs from the Brazos River sites from shoal chubs 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 2).

Admixture proportions were calculated in Entropy for *k* = 2–4 with all of the sampled sites (Figure 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

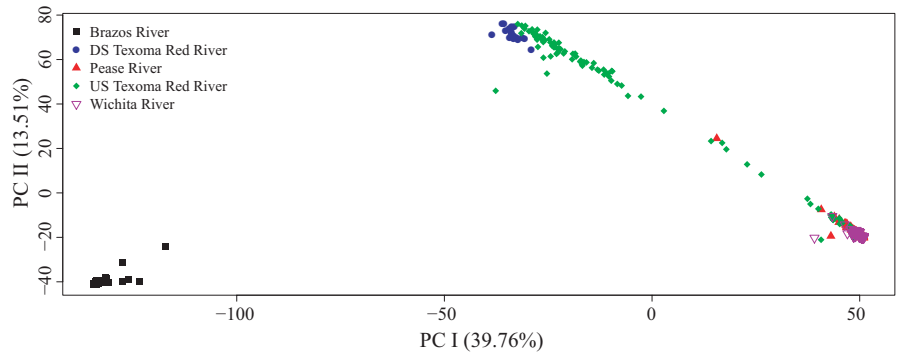

FIGURE 2 PCA of genetic differentiation of all individuals from each collection locale. PC I explains approximately 39.76% of the genetic variation in the data, and PC II explains approximately 13.51% of the variation. Shapes and colors represent different collection locales. Black squares are individuals 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

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.

3.3 | Genetic and environmental associations

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, pH, and current velocity; temperature was removed by the stepwise procedure. The selected model was significant ($F_{5,303} = 128.5$, $p < 0.001$) and explained approximately 67.43% of the variation in q (Table 3). Of the explained variation in the final model, specific conductance (95% CI: 34.36 < 40.08 < 45.80) explained 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 and depth.

3.4 | Admixture class

Admixture class estimates (Q_{12}) revealed mixed ancestry for individuals from several sites upstream from Lake Texoma (Figure 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

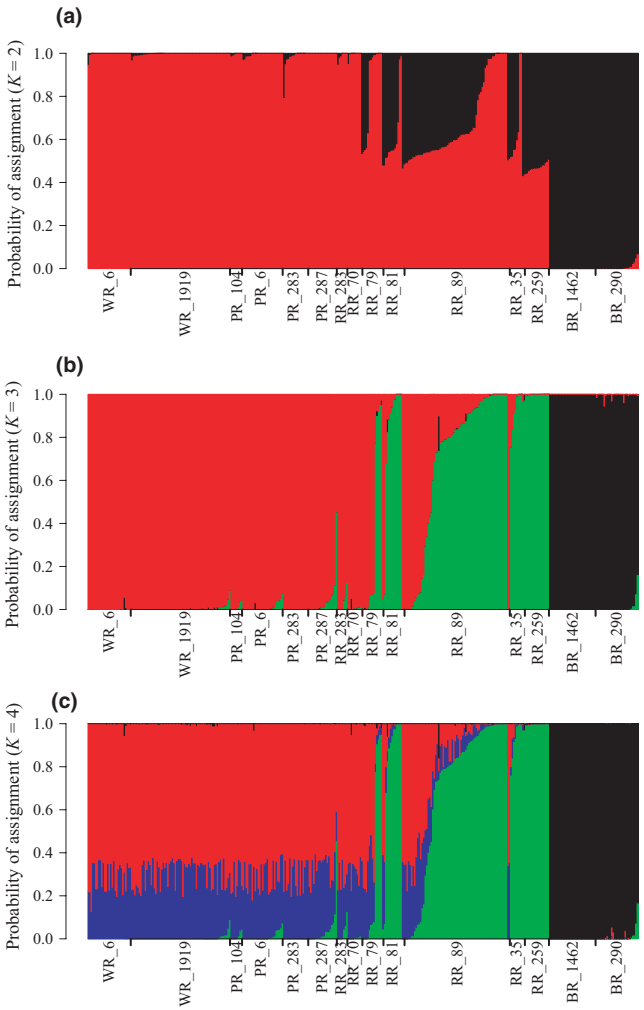

FIGURE 3 Entropy plots for $k = 2$ (a), 3 (b), and 4 (c) for all sampled sites oriented from most upstream to downstream (left to right) in the Red River basin, followed by the Brazos River. Site codes are defined in Table 1

TABLE 3 Output from multiple regression of environmental variables (specific conductance, water depth, dissolved oxygen, current velocity, and pH) predicting *q* from Entropy (assignment probability to first genetic cluster from Entropy), including estimated 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	0.538	8.793	<0.001	40.08	34.36	45.80
Depth	−0.287	−3.529	<0.001	34.83	31.04	38.16
Dissolved oxygen	−2.374	−7.023	<0.001	20.30	14.58	25.80
Current velocity	0.856	4.158	<0.001	3.87	2.76	5.62
pH	5.055	3.112	0.002	0.91	0.27	3.19

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 ancestry and 1.0 for interpopulation ancestry), indicating that while F_1 hybridization may be rare, those hybrids—as well as later-generation hybrids, are likely to be fit.

3.5 | Genomic clines

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 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 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 associated with reproductive isolation between the two taxa), and 18 had exceptionally negative β values (indicating that bidirectional selective introgression is likely to occur for these loci). We 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 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.

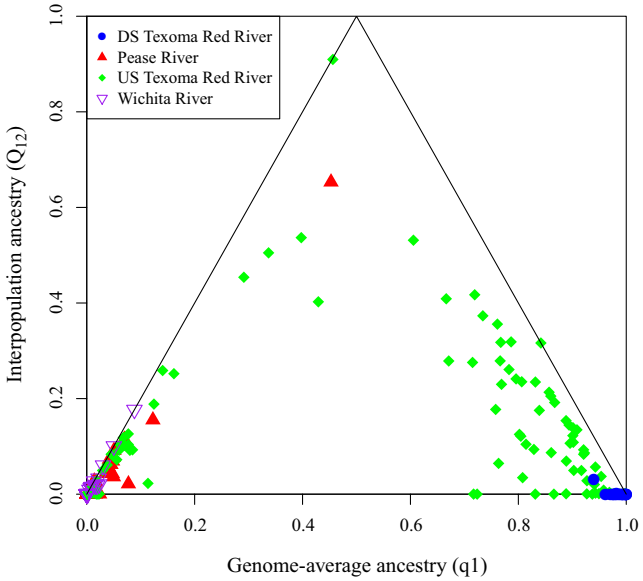

FIGURE 4 Scatter plot showing the relationship between genome-average ancestry (q1) and interpopulation ancestry (Q_{12}). Symbols correspond to individuals from areas of the Red River basin (DS Texoma Red River = RR_259; Pease River = PR_287, PR_283, PR_6, and PR_104; US Texoma Red River = RR_35, RR_89, RR_81, RR_79, RR_70 and RR_283; Wichita River = WR_1919 and WR_6). Solid lines indicate the maximum possible interpopulation ancestry given global genetic ancestry

4 | DISCUSSION

4.1 | Genetic structuring and hybridization

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 (Echelle et al., 2018; Eisenhour, 2004; Underwood et al., 2003). 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 and trace amounts of admixture evident in the Wichita River. As in previous studies, we found higher genetic differentiation between two populations of shoal chubs in different drainages (Brazos River and

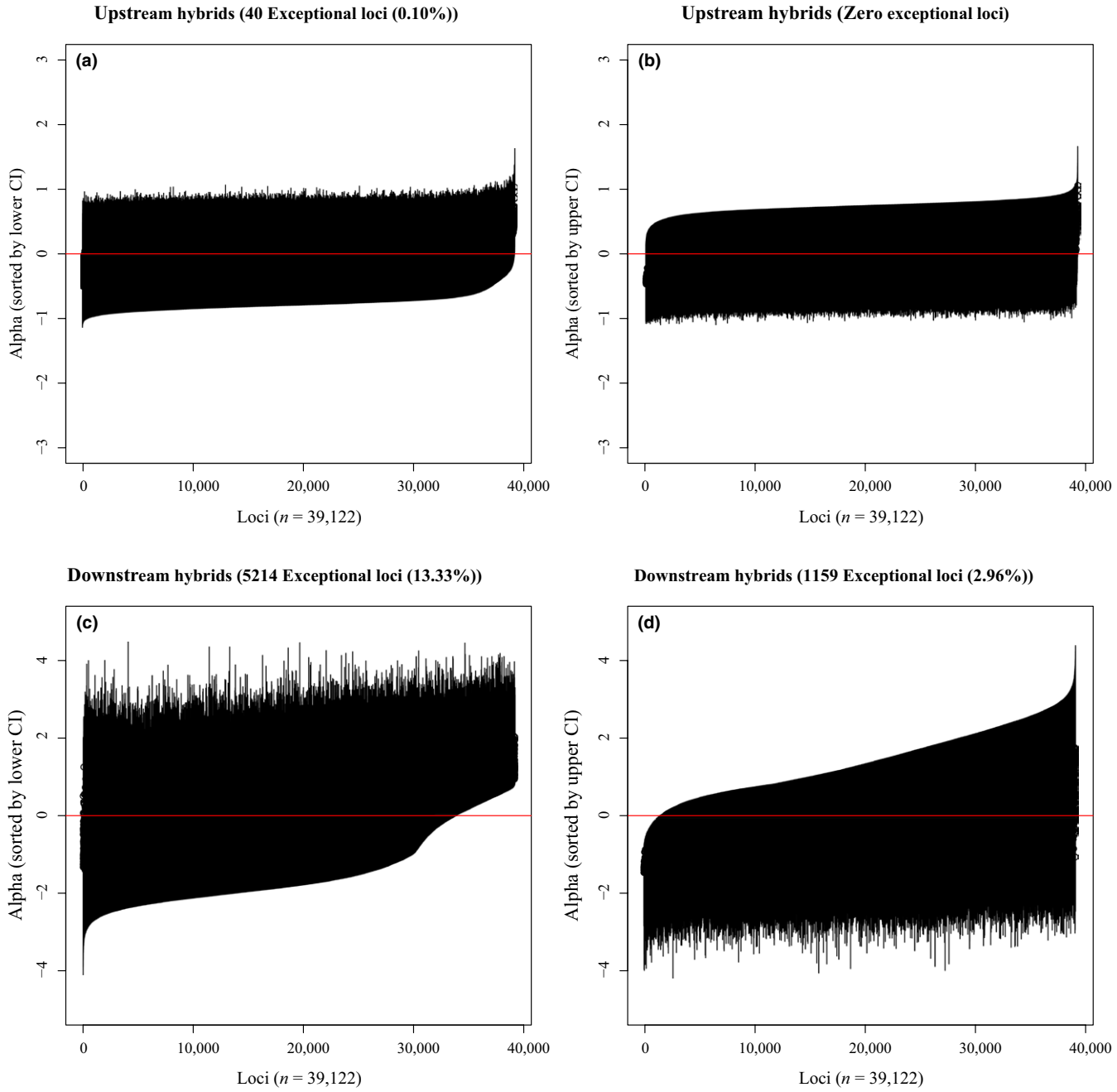

FIGURE 5 Median (± 95% CIs) of Bayesian genomic cline parameter α sorted by the lower 95% CI and by the upper 95% CI for hybrid individuals captured in the upstream (a, b) and downstream Red River (c, d). Upstream hybrids include individuals captured at RR_283, PR_287, PR_6, and WR_1919; downstream hybrids 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

Red River) than between shoal chubs and prairie chubs within the Red River (Echelle et al., 2018; Underwood et al., 2003). We 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 in the hybrid zone, with hybrid individuals consisting predominately of shoal chub background, and introgression of alleles, with shoal chub alleles introgressing at a higher frequency than prairie chub alleles. Attempts to classify the hybrid individuals into early-generation hybrid classes (e.g., F_1, F_2, or BC_1 individuals) were largely unsuccessful, with only a single mixed-ancestry individual being categorized as a possible F_1. 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 in the hybrid zone, 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. Asymmetric hybridization been attributed to several factors including differences in generation time (Barton, 1986), mating behavior (Konkle & Philipp, 1992; Lamp & Avise, 1986), fitness (Ostberg, Slatton, & Rodriguez, 2004), or relative abundances of the parental species (Lepais et al., 2009). 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 and hybrid individuals having more mating opportunities with heterospecifics than with conspecifics. Greater abundance of shoal chubs and lesser abundance of prairie chubs within the hybrid zone corresponds with an overlap in distributions between prairie chubs and shoal chubs indicating this hybrid zone might be a natural secondary contact zone between the two species and therefore represent natural hybridization 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, Gephard, & Pescitelli, 2005), which can lead to an increase in introgressive hybridization (Hasselman et al., 2014; Seehausen, Takimoto, Roy, & Jokela, 2008). Thus, it is possible that the construction of Lake Texoma could have restricted a species of mobile *Macrhybopsis* (Wilde, 2016; Worthington et al., 2016) upstream and, along with altering upstream habitats (e.g., deeper water, more similar to habitats associated with shoal chubs; Eisenhour, 2004), could have anthropogenically inflated shoal chub numbers in the zone.

Bayesian genomic cline 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., 2014a, 2014b, 2012; Kingston, Parchman, Gompert, Buerkle, & Braun, 2017; Parchman et al., 2013; Payseur, 2010; Teeter et al., 2010; Yuri, Jernigan, Brumfield, Bhagabati, & Braun, 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, Randell, & Rieseber, 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) suggest 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 (Aboim, Mavarez, Bernatchez, & Coelho, 2010; Nolte et al., 2009). 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.

4.2 | Genetic and environmental associations

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 et al. (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, we 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. We found that temperature was not a significant predictor of q, which is in contrast to other published studies of fish hybridization (Marie et al., 2012; Yau & Taylor, 2013). 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 the distributions of these species. 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 (Higgins & Wilde, 2005; Ruppel et al., 2017). Additionally, there are no physical barriers preventing the prairie chub from moving lower in the basin toward 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 importance of water quality in shaping aquatic fish communities (Barlow, 1958; Marie et al., 2012; Ostrand, 2000; Ostrand & Wilde, 2001; 2004), this certainly presents an interesting hypothesis and warrants future experiments or studies to understand species restrictions.

5 | CONCLUSIONS

Overall, we 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. There is also a broad pattern of 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 (Aboim et al., 2010; Nolte et al., 2009), which is not the case here. We found that assignment probability was predicted by various water quality parameters (e.g. specific conducatnce, depth, dissolved oxygen, current velocity, and pH), indicating that the location of these species and their hybrids is highly associated with water quality. Thus, if the water quality was 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), we 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 (Echelle et al., 2018; Eisenhour, 2004; 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.

ACKNOWLEDGMENTS

Thank you to all the graduate and undergraduate students at Texas State University which assisted with this research, specifically Cody Craig, Jeremy Maikoetter, Nicky Hahn, Alex Zalmat, and Sarah Bialik. This research was funded by the Texas Comptroller of Public Accounts [RFP No. 212f].

AUTHORS' CONTRIBUTIONS

V.A.S., T.H.B., and N.H.M designed the research; V.A.S., T.H.B., and D.S.R. conducted the field work; V.A.S. conducted the laboratory work; V.A.S., C.C.N., and N.H.M. analyzed the data; V.A.S., D.S.R., C.C.N., T.H.B., and N.H.M. wrote the manuscript.

DATA ACCESSIBILITY

Data (individual fastq files containing sequence reads, and environmental data from sampling sites) are currently accessible from the Dryad Digital Repository: https://doi.org/10.5061/dryad.dr06p8f.

ORCID

V. Alex Sotola https://orcid.org/0000-0002-9304-1017
Noland H. Martin https://orcid.org/0000-0001-7646-3501

REFERENCES

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. https://doi.org/10.1111/j.1420-9101.2010.01953.x

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. https://doi.org/10.1016/S0169-5347(01)02290-X

Arnold, M. L., & Martin, N. H. (2009). Adaptation by introgression. *Journal of Biology*, *8*, 82. https://doi.org/10.1186/jbiol176

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: Biological Sciences*, *363*, 2813–2820. https://doi.org/10.1098/rstb.2008.0021

Barlow, G. W. (1958). High salinity mortality of desert pupfish, *Cyprinodon macularius*. *Copeia*, *1958*, 231–232. https://doi.org/10.2307/1440605

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

Buerkle, C. A., & Lexer, C. (2008). Admixture as the basis for genetic mapping. *Trends in Ecology and Evolution*, *23*, 686–694. https://doi.org/10.1016/j.tree.2008.07.008

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

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

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. https://doi.org/10.1111/j.1365-294X.2006.03148.x

Eisenhour, D. J. (2004). Systematics, variation, and speciation of the *Macrhybopsis aestivalis* complex west of the Mississippi River. *Bulletin Alabama Museum of Natural History*, *23*, 9–48.

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.

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. https://doi.org/10.1093/bioinformatics/bts565

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. https://doi.org/10.1111/j.1365-294X.2011.05074.x

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 challeneges 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. (2014a). Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology*, *23*, 4555–4573.

Gompert, Z., Lucas, L. K., Buerkle, C. A., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2014b). Data from: Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.pq93h

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. https://doi.org/10.1111/j.1558-5646.2012.01587.x

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

Groemping, U. (2006). Relative importance for linear regression in R: The package relaimpo. *Journal of Statistical Software*, *17*, 1–27.

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

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. https://doi.org/10.1111/mec.12674

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. https://doi.org/10.1111/j.1752-4571.2009.00100.x

Higgins, C. L., & Wilde, G. R. (2005). The role of salinity in structuring fish assemblages in a prairie stream system. *Hydrobiologia*, *549*, 197–203. https://doi.org/10.1007/s10750-005-0844-7

Hubbs, C. L. (1955). Hybridization between fish species in nature. *Systematic Zoology*, *4*, 1–20. https://doi.org/10.2307/2411933

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

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. https://doi.org/10.1111/jeb.13033

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

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 of the United States of America*, *83*, 2526–2530.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

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. https://doi.org/10.1111/j.1365-294X.2009.04137.x

Levin, D. A. (1983). Polyploidy and novelty in flowering plants. *The American Naturalist*, *122*, 1–25. https://doi.org/10.1086/284115

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

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. https://doi.org/10.1093/bioinformatics/btr509

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 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. https://doi.org/10.1093/bioinformatics/btl158

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. (2005). Loci affecting long-term hybrid survivorship in Louisiana irises: Implications for reproductive isolation and introgression. *Evolution*, *59*, 2116–2124. https://doi.org/10.1111/j.0014-3820.2005.tb00922.x

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. https://doi.org/10.1534/genetics.105.053538

Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols*, *2010*(6), pdb-prot5448.

Nadeau, N. (2014). Butterfly genomics sheds light on the process of hybrid speciation. *Molecular Ecology*, *23*, 4441–4443. https://doi.org/10.1111/mec.12877

NatureServe (2010). Digital Distribution Maps of the Freshwater Fishes in the Conterminous United States. Version 3.0. Arlington, VA

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

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. https://doi.org/10.1111/j.1365-294X.2009.04208.x

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. https://doi.org/10.1111/j.1365-294X.2004.02268.x

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.

Ostrand, K. G., & Wilde, G. R. (2004). Changes in prairie stream fish assemblages restricted to isolated streambed pools. *Transactions of the American Fisheries Society*, *133*, 1329–1338.

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*, 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., … Buerkle, C. A. (2013). The genomic consequences of adaptive divergence and reproductive isolation between species of manakins. *Molecular Ecology*, *22*, 3304–3317. https://doi.org/10.1111/mec.12201

Payseur, B. A. (2010). Using differential introgression in hybrid zones to identify genomic regions involved in speciation. *Molecular Ecology Resources*, *10*, 806–820. https://doi.org/10.1111/j.1755-0998.2010.02883.x

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. https://doi.org/10.1002/aqc.2501

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. https://doi.org/10.7717/peerj.431

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, *27*, 83–109. https://doi.org/10.1146/annurev.ecolsys.27.1.83

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.

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.

Santucci, V. J. Jr, 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.

Seehausen, O., Takimoto, G., Roy, D., & Jokela, J. (2008). Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology*, *17*, 30–44.

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, *60*, 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

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.

Teeter, K. C., Thibodeau, L. M., Gompert, Z., Buerkle, C. A., Nachman, M. W., & 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., Eisenhour, D. J., Jones, M. D., Echelle, 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*, *3*, 493–501. https://doi.org/10.1643/CG-02-033R1

Whitney, K. D., Randell, R. A., & Rieseber, L. H. (2006). Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *The 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.

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 of the United States of America*, *106*, 13875–13979. https://doi.org/10.1073/pnas.0811575106

Worthington, T. A., Echelle, A. A., Perkin, J. S., Mollenhauer, R., Farless, N., Dyer, J. J., … 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. https://doi.org/10.1046/j.1420-9101.2001.00335.x

Yau, M. M., & Taylor, E. B. (2013). Environmental and anthropogenic correlates of hybridization between westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and introduced rainbow trout (*O. mykiss*). *Conservation Genetics*, *14*, 885–900. https://doi.org/10.1007/s10592-013-0485-8

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.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Sotola VA, Ruppel DS, Bonner TH, Nice CC, Martin NH. Asymmetric introgression between fishes in the Red River basin of Texas is associated with variation in water quality. *Ecol Evol*. 2019;9:2083–2095. https://doi.org/10.1002/ece3.4901