

9-2019

Late Pleistocene Range Expansion of North American Topminnows Accompanied by Admixture and Introgression

David D. Duvernell

Missouri University of Science and Technology, duvernell@umsl.edu

Eric Westhafer

Southern Illinois University

Jacob F. Schaefer

University of Southern Mississippi, Jake.Schaefer@usm.edu

Follow this and additional works at: https://aquila.usm.edu/fac_pubs



Part of the [Biology Commons](#)

Recommended Citation

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*, 46(9), 2126-2140.

Available at: https://aquila.usm.edu/fac_pubs/16553

Late Pleistocene range expansion of North American topminnows accompanied by admixture and introgression.

Running Title: Pleistocene range expansion of topminnows

David D. Duvernell^{*1,2}, Eric Westhafer² and Jacob F. Schaefer³

¹ Missouri University of Science and Technology

² Southern Illinois University Edwardsville

³ University of Southern Mississippi

* Corresponding author – duverneld@mst.edu

Acknowledgements – We thank S. Clark, L. Stewart, B. Kreiser, J. Harris, and T. McGowan for assistance with field collections. Lab assistance with DNA sample preparation was provided by A. Cracchiolo, H. Nguyen, L. Macke and C. Watson. Samples were collected under the authority of collector permits issued by relevant state agencies. Funding for this project was provided by NSF DEB-1556778 as well as resources provided by SIUE and USM.

Abstract

Aim

We used genome-scale sampling to assess the phylogeography of a group of topminnows in the *Fundulus notatus* species complex. Two of the species have undergone extensive range expansions resulting in broadly overlapping distributions, and sympatry within drainages has provided opportunities for hybridization and introgression. We assessed the timing and pattern of range expansion in the context of late-Pleistocene-Holocene drainage events, and evaluated the evidence for introgressive hybridization between species.

Location

Central and southern United States including drainages of the Gulf of Mexico Coastal Plain and portions of the Mississippi River drainage in and around the Central Highlands.

Taxon

Topminnows, Genus *Fundulus*, subgenus *Zygonectes* - *Fundulus notatus*, *Fundulus olivaceus*, *Fundulus euryzonus*

Methods

We sampled members of the *F. notatus* species complex throughout their respective ranges, including numerous drainage systems where species co-occur. We collected genome-wide Single Nucleotide Polymorphisms (SNPs) using the Genotype by Sequencing (GBS) method, and subjected data to population genetic analyses to infer the population histories of both species, including explicit tests for admixture and introgression. The methods employed included STRUCTURE, Principle Coordinates Analysis, TreeMix and Approximate Bayesian Computation.

Results

Genetic data are presented for 749 individuals sampled from 14 *F. notatus*, 20 *F. olivaceus*, and 2 *F. euryzonus* populations. Members of the species complex differed in phylogeographic structure, with *F. notatus* exhibiting geographic clusters corresponding to Pleistocene coastal drainages, and *F. olivaceus* comparatively lacking in phylogeographic structure. Evidence for interspecific introgression varied by drainage.

Main Conclusions

Populations of *F. notatus* and *F. olivaceus* exhibited contrasting patterns of lineage diversity among coastal drainages, indicating inter-specific differences in their Pleistocene southern refugia. Phylogeographic patterns in both species indicated that range expansions into the northern limits of contemporary distributions coincided with and continued subsequent to the Last Glacial Maximum. There was evidence of introgression between species in some, but not all drainages where the species co-occur, in a pattern that is correlated with previous estimates of hybridization rates.

Keywords – Dispersal, Hybridization, Introgression, Admixture, Phylogeography, SNP, Population Genomics, Pleistocene, Central Highlands, Coastal Plain

Introduction

Global biodiversity patterns, including those in temperate freshwater systems, have been significantly impacted and shaped by climate variability associated with the Pleistocene-Holocene time period (Hewitt, 1999, 2004; Mittelbach et al., 2007). Large scale dynamic factors that influence freshwater biodiversity include fluvial geomorphology-altering events like sea-level fluctuation (Dias et al., 2014), and processes that alter historical drainage patterns (e.g. Jeffries et al., 2016; Kotlík, Bogutskaya, & Ekmekçi, 2004; Yokoyama, Sideleva, Shedko, & Goto, 2008; Carrea, Anderson, Craw, Waters, & Burrige, 2014). These factors have impacted how species were able to move and shift their distributions in response to changing environments (Gante et al., 2009; Gómez & Lunt, 2007; Hewitt, 1999, 2004; Soltis, Morris, McLachlan, Manos, & Soltis, 2006).

Much is known about the biogeographic history of the Central Highlands and Lowlands, and the Gulf Coastal Plain of North America through phylogeographic studies of the ichthyofauna (Bossu, Beaulieu, Ceas, & Near, 2013; Echelle et al., 2014; Egge & Hagbo, 2015; Hundt, Berendzen, & Simons, 2017; Near & Keck, 2005; Ray, Wood, & Simons, 2006; Schönhuth et al., 2018; Strange & Burr, 1997) and other aquatic species of the region (e.g. crayfishes (Crandall, Templeton, & Neigel, 1999), salamanders (Highton, 1995; Mayden, 1985), mussels (Zanatta & Harris, 2013)). The rich species diversity of the region is comprised of a complex assortment of endemics and widespread species (Cross, Mayden, & Stewart, 1986; Hoagstrom, Ung, & Taylor, 2014; Mayden, 1988). Many of the phylogenetic groups exhibit allopatric distributions shaped by vicariance and dispersal events (Bossu et al., 2013; Wiley & Mayden, 1985) with evidence of hybridization and introgression often associated with secondary contact (Dowling, Broughton, & DeMarais, 1997; Harrington, Benavides, & Near, 2013; Wallis et al., 2017). Along the Gulf Coastal Plain, the role of climate-driven sea-level fluctuations on coastal drainage connectivity has had a significant influence on phylogeographic patterns and distributions exhibited by coastal drainage freshwater species (Dias et al., 2014; Mayden & Allen, 2015; Swift, Gilbert, Bortone, Burgess, & Yerger, 1986).

The *Fundulus notatus* species complex (*F. notatus*, *F. olivaceus*, *F. euryzonus*) is a group of closely related topminnow species (Thomerson, 1966). The distinctive feature of this group is its broad distribution across drainage systems, as well as widespread sympatry of member species (Duvernell, Meier, Schaefer, & Kreiser, 2013; Wiley & Mayden, 1985), which provides a valuable perspective to our understanding of North American freshwater Pleistocene-Holocene biogeography (Duvernell et al., 2013). The broad sympatry of member species also serves as a model for understanding speciation, diversification, and reproductive isolation through numerous contact zones replicated at the scale of isolated drainage systems (Duvernell & Schaefer, 2014; Schaefer, Duvernell, & Campbell, 2016).

The center of diversity for the *F. notatus* species group lies within coastal drainages in Mississippi and Louisiana, where all three species can be found in sympatry (Howell & Black, 1981; J. Schaefer, Kreiser, Champagne, Mickle, & Duvernell, 2009) (Fig. 1). *Fundulus notatus* populations have diverged into four phylogenetic clades, distributed along the Gulf Coastal Plain (Duvernell et al., 2013). The ranges of each of the respective clades correspond to historic Pleistocene drainages (Galloway, Whiteaker, & Ganey-Curry, 2011; Snedden et al., 2018). In contrast, *F. olivaceus* exhibits very limited geographic structure across its range, which is largely overlapping with *F. notatus*. The third species, *F. euryzonus*, is restricted to populations in drainages where all three species co-occur, including the Amite River and Pascagoula River drainages in Louisiana and Mississippi.

Members of the *Fundulus notatus* group are habitat generalists and freshwater obligates with low salinity tolerance (Griffith, 1974). They are effective stream dispersers with the ability to occupy niches throughout stream gradients from headwaters to margins and backwaters of large rivers. In allopatry, both *F. notatus* and *F. olivaceus* can be found throughout this gradient (Braasch & Smith, 1965; Thomerson, 1966). Paradoxically, in drainages where both species co-occur, they tend to sort along stream ecological gradients. Areas where the species are syntopic tend to be limited spatially, most often near confluences (Duvernell & Schaefer, 2014). The pattern of which species occupies up vs. downstream habitat is inconsistent among drainages, and may reflect a combination of local ecological factors as well as colonization history (Duvernell & Schaefer, 2014; Schaefer et al., 2016).

Recent latitudinal range expansion has been offered as an explanation for limited range-wide nuclear and mitochondrial phylogeographic structure for *F. olivaceus*, and similarly limited geographic structure for *F. notatus* outside of defined gulf coastal drainages (Duvernell & Schaefer, 2014). In this study, we have employed genome-wide sampling of single nucleotide polymorphisms (SNPs) to elucidate high resolution inference of population historical relationships, and infer the timing of range expansion of the respective species. We evaluate the hypothesis that all three topminnow species occupied Pleistocene coastal drainage refugia, and that range expansions coincided approximately with the Last Glacial Maximum (LGM), determining contemporary distributions of both *F. notatus* and *F. olivaceus*. We also evaluate the extent of admixture among species lineages, and ascertain if secondary contact has resulted in introgression between sympatric populations of *F. notatus* and *F. olivaceus*.

Methods

We sampled populations of *F. notatus*, *F. olivaceus* and *F. euryzonus* from throughout their respective ranges, capturing all major known phylogenetic lineages (Duvernell et al., 2013). Most specimens were collected by dipnet at public road access points. Fin clips were preserved in either 100% ethanol or a high salt preservative (Seutin, White, & Boag, 1991), and genomic DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit (Qiagen). We employed Genotype by Sequencing (GBS) to obtain Single Nucleotide Polymorphism (SNP) genotypes following Elshire et al. (2011). The EcoT221 restriction enzyme was used to construct GBS libraries, and sequencing was performed using Illumina HiSeq. Raw sequence data, consisting of barcoded 64 base reads, was analyzed with the TASSEL (version 4.0) GBS pipeline (Bradbury et al., 2007; Glaubitz et al., 2014) using a *F. heteroclitus* reference genome (<https://my.mdibl.org/diplay/FGP/Home>) as previously reported (Schaefer et al., 2016). Sequences were aligned to the *F. heteroclitus* genome using bowtie 2.0 (Langmead & Salzberg, 2012) using the default “very-sensitive” end-to end option, yielding an alignment rate of 38%. Heterozygote calls were made with a quantitative SNP calling function utilizing read counts and an expected sequencing error rate of 0.01 (Bradbury et al., 2007). We ran TASSEL three times to export raw SNP genotypes for 1) all three species plus outgroups, 2) *F. olivaceus* alone, and 3) *F. notatus* alone.

Raw SNP genotypes were imported into R for filtering of data prior to construction of input files for various analyses using custom R scripts. Data files included only bi-allelic loci filtered on missing data by locus, by individual, and by minimum total allele frequency. Filter thresholds varied by analysis and are reported in supplementary information (Appendix 1, Table S1.1). A small subset (~6%) of loci exhibited substantial excess observed heterozygosity (approaching 100%) that likely represents artifacts of misalignment of paralogs (Drury et al., 2016; Nunez et al., 2015). We filtered for this artifact by conservatively removing all loci exhibiting > 50% heterozygosity. Finally, loci were filtered for locus

proximity to minimize linkage disequilibrium. We filtered to a minimum distance of 5000 bp for loci sharing alignment to the same scaffold in the reference genome.

Some samples were previously reported in a study of hybridization and introgression in topminnow contact zones (Schaefer et al., 2016), and other samples have been collected for the purpose of studying additional contact zones. We avoided inclusion of recent-generation hybrids and backcross individuals in this study by including only collection sites that were sampled outside known boundaries of syntopy within sympatric drainages. We performed an initial screen of all *F. notatus* and *F. olivaceus* collections to assess admixture status by conducting a STRUCTURE analysis (v. 2.3.4; (Pritchard, Stephens, & Donnelly, 2000), with $K = 2$ (average of five replicates of one million repetitions after a burnin of 100,000). In this analysis K corresponds to the number of populations, or separate gene pools, set by the investigator (Pritchard et al., 2000). Individuals with coefficient of membership (Q -score) ≤ 0.969 (equivalent to 4th generation or more recent backcross) were excluded from all subsequent analyses.

Divergence and Intraspecific Population Genetic Structure

We explored the range-wide and regional population genetic structure within each species using a STRUCTURE analysis as well as a Principle Coordinates Analysis (PCoA). We adjusted sample sizes to ten individuals per location by rarefaction, when necessary, to achieve similar sample sizes across locations. The analysis parameters included variable K , with one million repetitions after a burnin of 100,000. Each analysis was replicated five times at each level of K , separate runs were combined using Structure Harvester web v. 0.6.94 (Earl & vonHoldt, 2012), and the optimal K selected using the Evanno method (Evanno, Regnaut, & Goudet, 2005). The PCoA analysis was conducted on a Euclidean similarity matrix (after converting SNPs to a 012 matrix) using the adegenet package in the R environment. Estimates of F_{ST} were obtained using Arlequin v. 3.5 (Excoffier & Lischer, 2010).

Coalescent Analysis of SNP Phylogeny

We assessed phylogenetic relationships of member species and regional population clusters of the *F. notatus* complex with SNAPP v. 1.1.1 (Bryant, Bouckaert, Felsenstein, Rosenberg, & RoyChoudhury, 2012) implemented in BEAST v. 2.3.1 (Drummond & Rambaut, 2007). This method estimates the genealogies of a set of unlinked bi-allelic SNP markers based on the coalescent. Uncertainty in phylogenetic relationships inferred by this method does not differentiate between incomplete lineage sorting concordant with recent range expansion, or introgression facilitated by secondary contact. The computational requirements of this analysis necessitated limiting each population cluster to a pool of five to seven individuals. Selection of individuals for membership in each cluster was determined using the results of the STRUCTURE analysis (above). Outgroups included individuals of *F. notti* and *F. chrysotus*. The single Markov chain Monte Carlo (MCMC) chain length was set to 1 million, with sampling every 1000 iterations, burnin set to 100,000, and using default priors and model parameters. Acceptable mixing was confirmed by inspection using Tracer v. 1.7 (Drummond & Rambaut, 2007), with effective sample size values >200 for all parameters. The distribution of trees was visualized using DENSITREE v. 2.1 (Bouckaert, 2010), and the maximum clade credibility tree was generated using TREEANNOTATOR v. 1.7.4 (Drummond & Rambaut, 2007).

Inference of Secondary Contact and Admixture Events

To visualize bifurcating population relationships, and infer the presence of admixture resulting from migration events among populations within and between species, we used the maximum likelihood approach implemented in TreeMix v. 1.12 (Pickrell & Pritchard, 2012). This method provides an assessment of whether shared variation among populations is better explained by incomplete lineage sorting, or secondary contact and admixture. TreeMix employs a graph-based model to represent relationships between populations, in the form of population splits and gene flow events. The weights of admixture events (w) provide an estimate of the proportion of ancestry of admixed populations. The fit of the data to the bifurcating model with admixture events is represented by a map of the covariance matrix, where positive values indicate pairs of populations that are more similar than the model represents, while negative values indicate population pairs that are less similar than indicated (Pickrell & Pritchard, 2012).

Estimation of the Timing of Population Range Expansion

We used Approximate Bayesian Computation (ABC), implemented in DIYABC v. 2.1 (Cornuet et al., 2014), to estimate the timing of range expansion for *F. notatus* and *F. olivaceus* from the posterior distribution of parameters estimated for select historical range expansion and admixture scenarios. ABC (Beaumont, 2010; Csilléry, Blum, Gaggiotti, & François, 2010; Estoup et al., 2012) utilizes simulations to generate large datasets drawn from prior parameter distributions (e.g. population sizes, coalescent times, admixture proportions) for given evolutionary scenarios. Then population statistics (for diversity, divergence and admixture) are used to summarize the simulated datasets and the actual population data, and Principle Components Analysis is used to find the simulated dataset parameter values that best match the actual population data.

Evolutionary scenarios were selected using inferences of population history provided by the TreeMix, and STRUCTURE analyses. We used relatively simple scenarios requiring only four or five population samples, to capture estimates of the timing of range expansion for each species. We also incorporated one intra-specific admixture event in each analysis (Appendix 2, Figure S2.1). Sample sizes ranged from 10 to 20 individuals per population. With SNP data (unlike with DNA sequence and microsatellite data), the DIYABC algorithm does not incorporate a mutation rate parameter. Therefore, the algorithm attempts to simultaneously fit population size and coalescence time parameters to the data. When both sets of parameters are unknown, the selection of prior distributions (Appendix 1, Table S1.2, S1.3) may have a significant impact on the outcome, and the resulting fitted time parameters should be viewed as scaled by effective population size, which is unknown. We evaluated use of a variety of summary statistics, including both F_{ST} and Nei distance metrics. Results that provided the best model fit to the data were reported. In the final analyses we conducted 1 million data simulations per evolutionary scenario, and compared scenarios by estimating posterior probabilities using the logistic regression (Cornuet, Ravigné, & Estoup, 2010). The goodness-of-fit test of the parameter posterior estimates using the model-checking option, and the assessment of bias and precision drawing on the posterior distributions, were both performed using default settings.

Results

Samples were collected from 14 *F. notatus*, 20 *F. olivaceus*, and 2 *F. euryzonus* localities (Fig. 1, Table 1), including all known major clades in the species group. Our dataset included 749 individuals after application of data quality filters. In an initial STRUCTURE analysis of *F. notatus* and *F. olivaceus*, we identified 25 individuals of putative hybrid ancestry with $Q < 0.969$, all in co-occurrence drainages, with

almost all other individuals exhibiting $Q > 0.992$ (more details in Appendix 2, Fig. S2.2). The median sample size, after removal of individuals with $Q < 0.969$, was 15 per location (Table 1). Genome SNP density and reference genome alignment details have been previously reported (Schaefer et al., 2016). Application of quality filters yielded data sets of 1000 to 3000 loci for most analyses, depending on which species, populations and individuals were included (See Appendix 1, Table S1.1).

Map	Stream	Drainage		Latitude	Longitude	<i>n</i>
Western Gulf Slope						
1	Colorado River	Colorado	<i>F. notatus</i>	29.7470	-96.5520	19
2	Angelina River	Neches	<i>F. notatus</i>	31.5855	-94.8293	9
3a	Sabine River	Sabine	<i>F. olivaceus</i>	30.6053	-93.7940	19
3b	Sabine River	Sabine	<i>F. notatus</i>	30.3073	-92.7247	18
Red River						
4	Glover River	Little/Red	<i>F. notatus</i>	34.2943	-94.9547	7
	Little River	Little/Red	<i>F. olivaceus</i>	33.9744	-94.9292	9
5	Mountain Fork	Little/Red	<i>F. notatus</i>	34.4617	-94.6352	2
6	Cossatot	Little/Red	<i>F. olivaceus</i>	33.8873	-94.2015	13
	Little River	Little/Red	<i>F. notatus</i>	33.7821	-94.1499	11
7	Kiamichi	Red	<i>F. olivaceus</i>	34.1985	-95.4846	15
			<i>F. notatus</i>	34.1992	-95.3503	8
8	Twelve Mile Bayou	Red	<i>F. notatus</i>	32.6456	-93.8774	12
9	Old River Bayou	Red	<i>F. olivaceus</i>	31.6975	-93.0751	10
Coastal Drainages						
10	Amite River	Amite	<i>F. euryzonus</i>	31.1284	-90.8485	7
11a	Pascagoula River	Pascagoula	<i>F. olivaceus</i>	30.7205	-88.7388	22
11b	Pascagoula River	Pascagoula	<i>F. notatus</i>	30.5662	-88.6261	22
			<i>F. euryzonus</i>	30.9873	-89.1948	8
12	Tombigbee River	Mobile	<i>F. olivaceus</i>	33.9853	-88.5511	28
			<i>F. notatus</i>	33.6623	-88.5538	63
13	Yellow River	Yellow	<i>F. olivaceus</i>	30.7047	-86.8814	20
Mississippi River						
14	Big Blue Hole	Mississippi	<i>F. notatus</i>	31.5823	-91.4823	16
15	Chotard Lake	Mississippi	<i>F. notatus</i>	32.5663	-91.0629	14
16	Elk River	Tennessee	<i>F. olivaceus</i>	34.8512	-87.2358	16
			<i>F. notatus</i>	34.8234	-87.2420	12
17	Duck River	Tennessee	<i>F. notatus</i>	35.7324	-86.9240	6
18	Spring River	Spring/Neosho	<i>F. notatus</i>	37.0258	-94.7239	11
	Shoal Creek	Spring/Neosho	<i>F. olivaceus</i>	37.0335	-94.5270	20
19	White River	White	<i>F. olivaceus</i>	36.9772	-92.1668	9
20	Black River	Black/White	<i>F. olivaceus</i>	37.4155	-90.8231	10
21	St. Francis River	St. Francis	<i>F. olivaceus</i>	37.6811	-90.4140	10
22	Cahokia Creek	Mississippi	<i>F. notatus</i>	38.8948	-89.9218	11
23a	Meramec River	Meramec	<i>F. olivaceus</i>	38.4173	-90.3475	37
23b	Meramec River	Meramec	<i>F. olivaceus</i>	38.4746	-90.6601	39
23c	Dry Fork	Meramec	<i>F. olivaceus</i>	37.8527	-91.6645	19
24	Big River	Meramec	<i>F. olivaceus</i>	37.7553	-90.8846	34
25a	Gasconade River	Gasconade	<i>F. olivaceus</i>	38.5533	-91.5974	35
25b	Gasconade River	Gasconade	<i>F. olivaceus</i>	37.9349	-91.9776	28
25c	Gasconade River	Gasconade	<i>F. olivaceus</i>	37.1947	-92.5616	33
26	Big Piney River	Gasconade	<i>F. olivaceus</i>	37.3331	-91.9582	43
	Calcasseiu River	Calcasseiu	<i>F. chrysotus</i>	30.6398	-92.8145	2
	Paul B. Johnson Res.	Pascagoula	<i>F. notti</i>	31.1537	-89.2423	1

Table 1: Collection locations and sample sizes for all population samples.

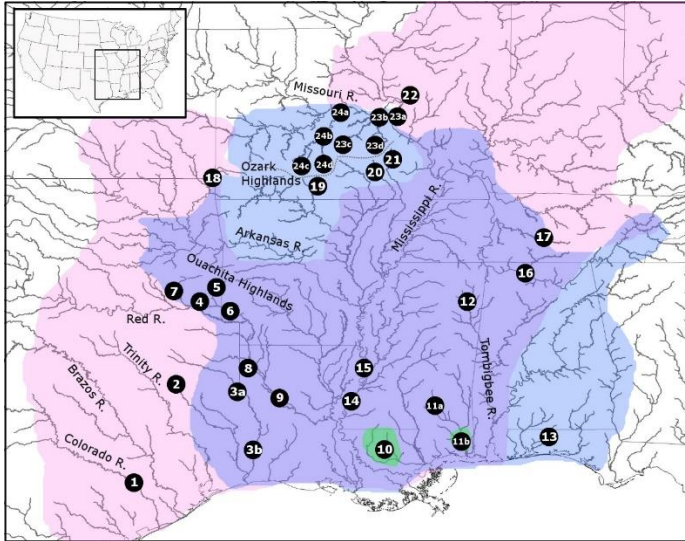


Figure 1: Species distributions for *F. notatus* (pink), *F. olivaceus* (blue) and *F. euryzonus* (green). The region of geographic overlap between *F. olivaceus* and *F. notatus* is shown in purple, and all three species are present in drainage regions shaded in green. Collection locations correspond to sample information provided in table 1.

A rooted Bayesian phylogenetic analysis of species and regional clusters within *F. notatus* and *F. olivaceus* supported the sister relationship between *F. olivaceus* and *F. euryzonus* (Fig. 2). The more defined regional structure of *F. notatus* relative to *F. olivaceus* was reflected in the deeper nodes. Multiple alternative regional relationships, visualized in the densitree for both *F. notatus* and *F. olivaceus* (Fig. 2), could reflect either secondary contact, or incomplete lineage sorting.

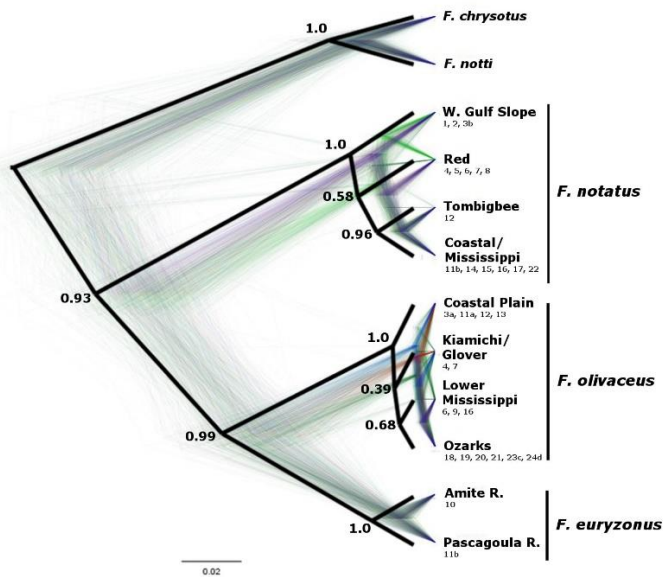


Figure 2: Phylogenetic relationships of member species and regional population groups based on a Bayesian analysis of SNPs. The samples represented in each regional group are numbered according to table 1. Numbers at nodes report posterior probability support for the consensus. The consensus phylogeny is superimposed on a densitree of alternative sampled trees, with contrasting topologies highlighted by different colors. The phylogeny is rooted with outgroup species, *F. notati* and *F. chrysotus*.

Estimates of F_{ST} among *F. notatus* populations ranged widely, from 0.01 to 0.9 (Appendix 1, Table S1.4). In *F. olivaceus*, estimates of F_{ST} ranged from 0.01 to 0.76 (Appendix 1, Table S1.5). Analysis of *F. notatus* populations supported four major clusters in both the STRUCTURE (Fig 3a) and PCoA (Fig. 3b) analyses. These clusters corresponded to the Western Gulf Slope (WGS), the Red River, the Mississippi River (plus coastal Pascagoula River), and the Tombigbee River. The only population sample that exhibited admixture was the Spring River, where individuals averaged 87.7% assignment to the Mississippi cluster, and 12.3% to the Red cluster.

Population structure was far more limited among *F. olivaceus* populations. A range-wide STRUCTURE analysis with $K = 2$ separated the Ozark Highlands and most of the Ouachita Highlands from coastal drainages (Fig. 4a). With $K = 3$, the Glover and Kiamichi Rivers within the Ouachita Highlands, but not the neighboring Cossatot River, were identified as genetically distinct. In the range-wide PCoA analysis of *F. olivaceus*, populations roughly sorted by latitude along the first axis, and Glover and Kiamichi Rivers separated along the second axis (Fig. 4c). A regional analysis of Ozark samples indicated two separate clusters (Fig. 4b, d).

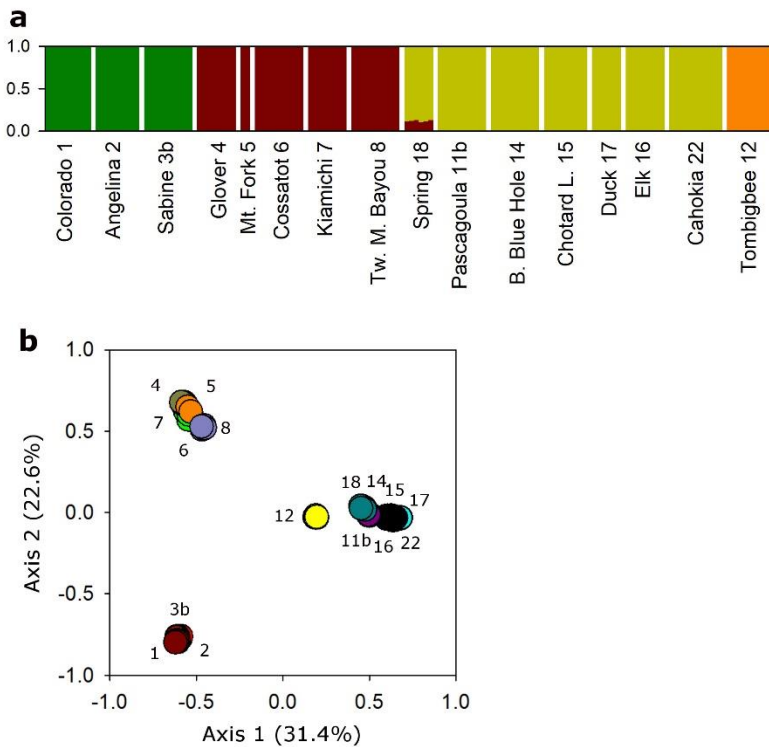


Figure 3: Analysis of *F. notatus* population relationships by a) STRUCTURE analysis ($K = 4$) and b) PCoA of individual genotypes. In the PCoA analysis, individuals are color coded according to their source population. Populations are numbered according to Fig. 1 and Table 1.

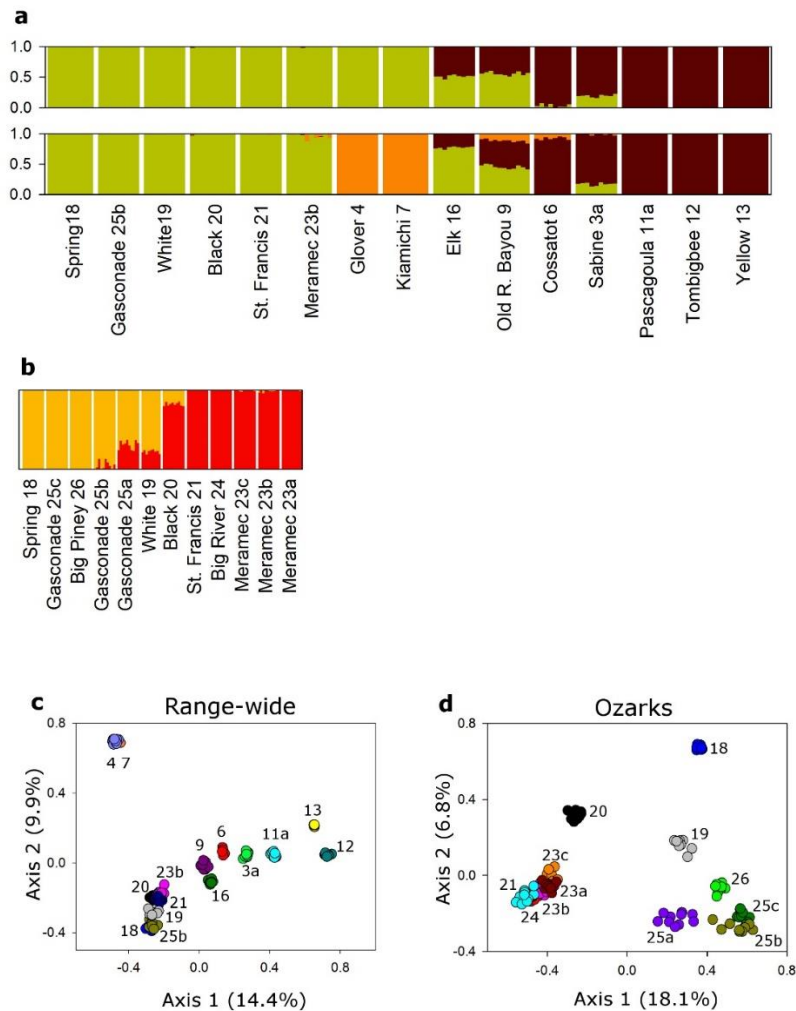


Figure 4: Hierarchical analysis of *F. olivaceus* population relationships revealed by STRUCTURE analysis of a) range-wide sampling ($K = 2, 3$), b) Ozark Highlands ($K = 2$), and c and d) PCoA analysis of the same populations. In the PCoA analysis, individuals are color coded according to their source population. Populations are numbered according to Fig. 1 and Table 1.

The maximum likelihood phylogeny constructed in TreeMix (Fig. 5) was concordant with the clustering relationships indicated by the STRUCTURE and PCoA analyses. The analysis inferred two intraspecific gene flow events among *F. notatus* clades. These included one originating near the base of the Kiamichi River branch, and terminating near the base of the Spring River branch ($w = 13\%$), consistent with admixture in the Spring River *F. notatus* population. A second intraspecific gene flow event ($w = 8.7\%$) originated at the base of the *F. notatus* WGS clade and extended to the base of the Red River clade.

Five interspecific gene flow events were inferred among the branch tips of the TreeMix phylogeny. These included *F. olivaceus* → *F. notatus* in the Tombigbee River ($w = 1.6\%$), Spring River ($w = 1.5\%$), and Kiamichi River ($w = 1.2\%$). The downstream population of *F. olivaceus* in the Little River near the confluence with the Glover River was connected via gene flow to the downstream population of *F. notatus* sampled at the confluence between the Cossatot and Little Rivers ($w = 1.4\%$). Finally, there was an inference of *F. notatus* → *F. olivaceus* within the Glover River ($w = 1.1\%$).

A TreeMix analysis of Ozark populations of *F. olivaceus* supported admixture ($w = 35\%$) of the lower Gasconade River (site 24a) from the Meramec River (site 23a) (Appendix 2, Fig. S2.3). Additional admixture events were also evident between north and south flowing headwater streams.

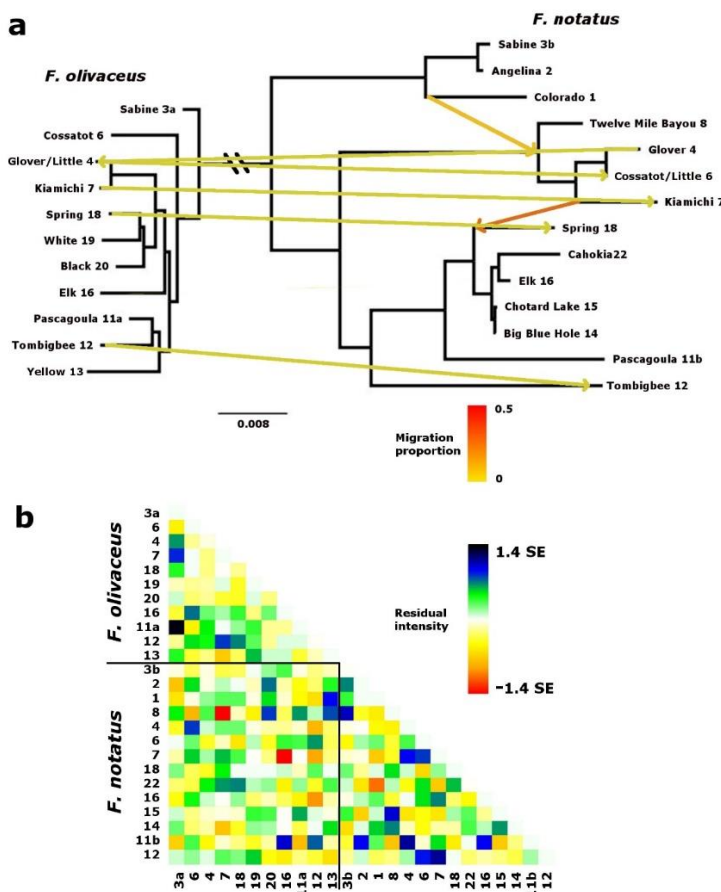


Figure 5: Maximum likelihood tree of *F. olivaceus* and *F. notatus* populations. a) population relationships of each species with migration events indicated by arrows color-coded according to their weight. b) A residual fit heat map with positive and negative residual values representing population pairs that are more or less closely related in the data than in the likelihood tree with migration.

We used ABC to evaluate the demographic histories and estimate the timing of population coalescence events of *F. notatus* and *F. olivaceus* populations in separate analyses. For our analysis of *F. notatus* (Fig. 6a), the Spring River admixture model was rejected in favor of a simpler bifurcating range expansion model (Appendix 2, Fig. S2.4). An assessment of the goodness-of-fit (model check analysis) demonstrated a marginal fit of the observed data to the posterior predictive distribution (Appendix 2, Fig. S2.5), a result that could not be improved with additional alternative models and population samples (data not shown).

In our evaluation of the demographic history of *F. olivaceus* (Fig. 6b), we focused on range expansion into the Ozark Highlands. The Gasconade River admixture model was strongly supported over either bifurcating model by both direct and logistic regression comparisons (Appendix 2, Fig. S2.6). The goodness-of-fit of the observed data to the admixture model was quite good (Appendix 2, Fig. S2.7).

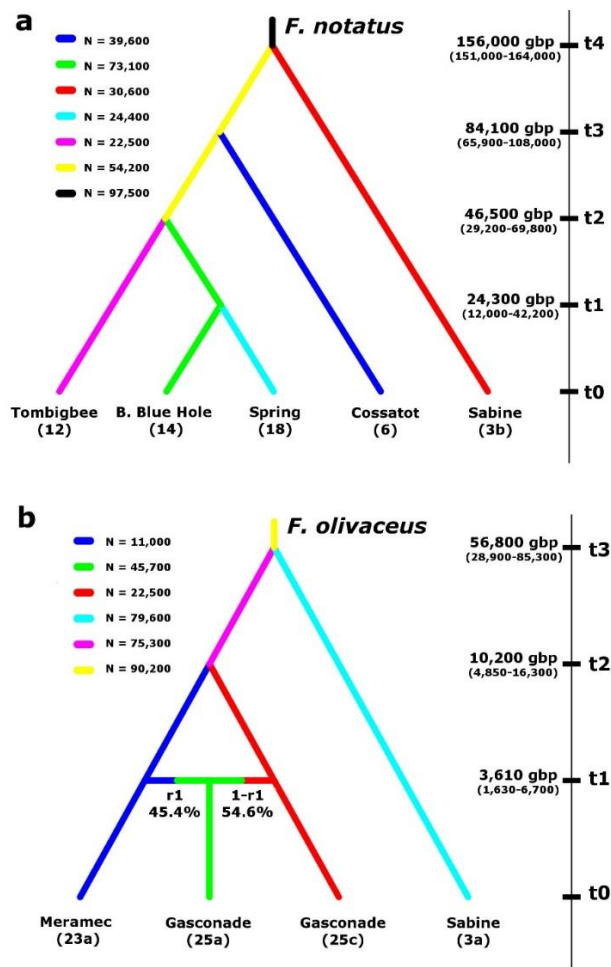


Figure 6: Approximate Bayesian Computation models for a) *F. notatus* and b) *F. olivaceus*. For each analysis, the model with the highest posterior probability is shown, along with estimates of effective population sizes (N), coalescence times in generations before present ($t \pm 95\%$ CI), and admixture proportions (r). Branch lengths do not scale with time.

Discussion

The broadly overlapping geographic ranges, and limited phylogeographic structure of *F. notatus* and *F. olivaceus*, suggest that both species have gone through recent parallel range expansions (Duvernell et al., 2013; Wiley & Mayden, 1985). Previous efforts to study evolutionary histories of the populations (e.g. Duvernell et al., 2013; Duvernell & Schaefer, 2014) have employed genetic markers (mitochondrial, AFLP, and a small number of nuclear loci) and sample sizes that were limited in their suitability for detecting and quantifying genome-wide admixture or interspecific introgression, or inferring the timing of population divergences associated with range expansion. In the present study, genome-wide SNP sampling provided high resolution genomic data for evaluating population relationships and inferring intra and interspecific genetic exchanges.

Phylogeographic history of the *F. notatus* species complex

A primary objective of this study was to investigate phylogeographic structure in *F. notatus* and *F. olivaceus*, and infer the timing of range expansion for each species, placing historical events in the context of known drainage relationships in the Gulf Coastal Plain and Interior Highlands regions. The distribution and diversity of fish faunal assemblages in the Gulf Coastal Plain has been attributed to freshwater drainage connections facilitated by low sea level events. Periods of high sea stands in the late Oligocene, Miocene and Pliocene have been associated with species-level diversification (Egge & Hagbo, 2015; Near, Kessler, Koppelman, Dillman, & Philipp, 2003; Swift et al., 1986), while sea level fluctuations during the Pleistocene have been implicated for population-level divergence (Dias et al., 2014; Swift et al., 1986). Drainages in the Gulf Coastal Plain served as important glacial refugia for a variety of species (Soltis et al., 2006), contributing to high species richness in the region (Oberdorff, Hugueny, & Guégan, 1997).

During the Pliocene and continuing through the Pleistocene, the central Gulf Coastal Plain and Interior Highlands were dominated by three major drainage systems corresponding to the Mississippi, Red, and Tennessee Rivers (Galloway et al., 2011; Snedden et al., 2018). Further to the east was the Mobile Basin, including the Tombigbee River, and to the west were a series of coastal drainages within the WGS (Hoagstrom et al., 2014). Frequent shifts in river courses were recorded in dispersed sediment fans along the gulf coast (Bentley, Blum, Maloney, Pond, & Paulsell, 2016), creating periodic drainage connections and dispersal opportunities for freshwater fauna (Hoagstrom et al., 2014). Only near the end of the Pleistocene the Red and Tennessee Rivers lost their independent connections to the gulf, and joined the Mississippi River Basin in their current configuration (Galloway et al., 2011).

Levels of interspecific mitochondrial divergence on the order of 6-9% (Duvernell et al., 2013) are consistent with species-level divergence that occurred during the Pliocene, or possibly earlier, for all three topminnow species. This study confirmed limited genetic structure among *F. olivaceus* populations relative to its contemporary geographic distribution. Our ABC modeling placed coalescence of all *F. olivaceus* populations at about 56,800 gbp, though with quite a bit of uncertainty (95% CI: 28,900-85,300 gbp). Assuming a one year generation time, this estimate is consistent with a single southern Pleistocene refugium.

This study confirmed our earlier observation (Duvernell et al., 2013) that *F. notatus* populations exhibit strikingly more range-wide geographic structure. *Fundulus notatus* populations cluster into four clades

that correspond broadly with major Pleistocene gulf coastal drainage systems (Fig. 7). Combining observations from this and a previous study (Duvernell et al., 2013), the four clades correspond to 1) the WGS clade, encompassing drainages that span the Texas coast between the Colorado and Sabine Rivers, as well as the upper reach of the Red River drainage (Duvernell et al., 2013). 2) The Red River clade, including the middle Red River drainage as well as tributaries originating in the Ouachita Highlands. 3) the Tombigbee River clade, in the Mobile basin. 4) the Mississippi River clade, including nearly the entire species distribution within the Mississippi River drainage (excluding the Red River), as well as gulf coastal drainages between the Mississippi and Tombigbee Rivers, including the Amite, Pearl and Pascagoula Rivers.

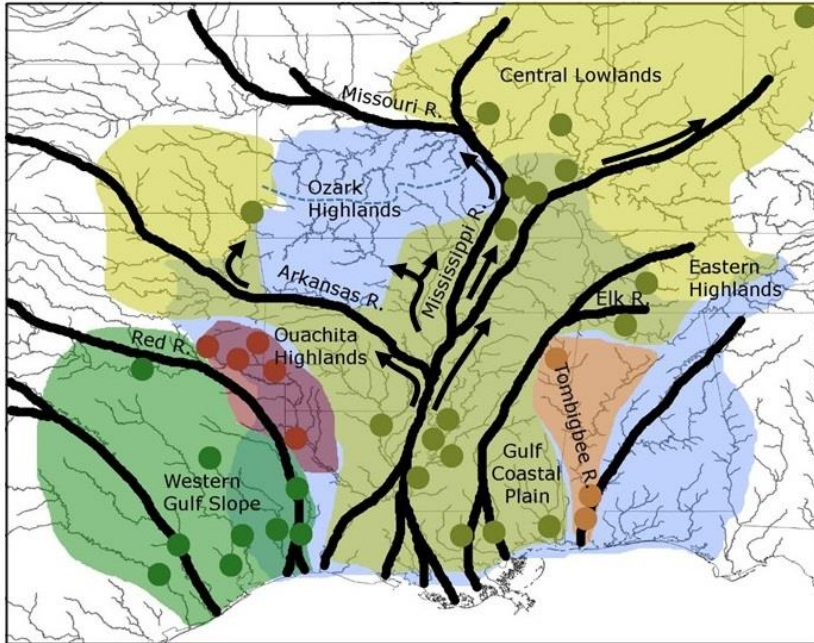


Figure 7: Map of the inferred distributions of four *F. notatus* clades corresponding to the WGS (green), Red River (red), Mississippi River and Gulf Coastal Plain (yellow), and Tombigbee River (orange). Supporting sample locations for *F. notatus* only, are drawn from the combined data presented in this study, and Duvernell et al. (2013). The distribution of *F. olivaceus* (blue) is also shown. Thick lines denote reconstructed late Pleistocene drainages modified from Galloway et al., (2011). Arrows indicate probable dispersal routes into the northern extents of ranges for both species. The Ozark Plateau divide between Missouri/Mississippi River tributaries and Arkansas/Black/St. Francis River drainages within the Ozark Highlands is indicated by a dotted line.

The coalescence times of *F. notatus* clades fit a pattern that is coincident with late Pleistocene climate cycles. The two most recent major low sea stands occurred during the Wisconsinan at the LGM, about 18,000 ybp, and during the Illinoian at 152,000 ybp, respectively, with additional minor fluctuations at more frequent intervals (Blum & Hattier-Womack, 2009; Imbrie et al., 1984; Lea, Martin, Pak, & Spero, 2002). According to genetic estimates in this study, and also assuming a generation time of one year, the earliest intraspecific coalescence event among *F. notatus* populations (95% CI: 151,000 to 164,000 gbp) was coincident with the Illinoian. Estimates of effective population size (N_e) in our ABC modeling efforts (Fig. 6) were substantially higher than those recently modeled in studies of other freshwater fish species by as much as 10-fold (e.g. Salisbury, McCracken, Keefe, Perry, & Ruzzante, 2016; Vera-Escalona, Habit, & Ruzzante, 2015). According to a Wright-Fisher model, effective population size and coalescent time

should scale proportionately. So, if our estimates of coalescent times were high, then our estimates of divergence times may be high as well. Given these considerations, we view coalescence time estimates reported here as upper limit estimates of actual values.

Genetic evidence indicated that the earliest split among *F. notatus* clades isolated the WGS group from other *F. notatus* populations. Subsequent to that initial event, admixture between the Red River and WGS clusters was suggested by contrasting *F. notatus* lineage genealogies represented in the densitree of our phylogenetic reconstruction of regional clusters (Fig. 2), and confirmed by gene flow at the base of the Red clade indicated in our TreeMix analysis (Fig. 5). Reconstructions of the Pleistocene course of the Red River indicate that it was at one time positioned further southwest of its current path, intersecting the present Sabine River drainage (Fig. 7). Sabine River *F. notatus* populations belong to the WGS clade (Fig. 2a). Other proposed connections between the Red River and WGS drainages include the hypothesized Pleistocene 'Plains Stream' (Cross et al., 1986; Metcalf, 1966) which would have connected portions of the upper Red River with the Trinity River. The Red River drainage has had a complicated history, and co-occurrence and admixture of the WGS and Red clades is consistent with this history (Duvernell et al., 2013). More comprehensive sampling of *F. notatus* populations in the Red River could provide further insights into the Pleistocene history of the drainage, and its association with the Ouachita Highlands and the WGS.

Our phylogenetic analysis of intraspecific *F. notatus* clade relationships (Fig. 2) improves on Duvernell et al. (2013) in which the relationships among the Red, Mississippi and Tombigbee clades were essentially unresolved. In this study, genome-wide SNPs indicated that, among the *F. notatus* populations, the Mississippi and Tombigbee clades split most recently, at about 46,500 gbp, with the Red clade becoming isolated longer ago, at about 84,100 gbp. Therefore, *F. notatus* clades are all of late Pleistocene origin. The most recent isolation between the Mississippi and Tombigbee clades is notable as it contradicts the most parsimonious explanation for chromosome evolution in this group of topminnows. Extensive sampling has demonstrated that the ancestral condition of $N = 24$ small acrocentric chromosomes is exhibited by *F. olivaceus* (Chen, 1971) and *F. euryzonus* (Howell & Black, 1981), while most *F. notatus* populations have a derived condition of $N = 20$ chromosomes (Chen, 1971). These karyotypes differ by the fusion of several smaller acrocentric chromosomes to form four larger metacentric Robertsonian chromosomes in *F. notatus* (Chen, 1971). Populations of *F. notatus* in the Tombigbee clade are distinct in exhibiting $N = 22$ chromosomes, with only two large metacentric chromosomes (Black & Howell, 1978). The most parsimonious explanation for chromosome evolution in this group would place the Tombigbee clade as basal to all other *F. notatus* clades, with an intermediate reduction in chromosome number between the ancestral (24) and derived (20) conditions. However, phylogenetic relationships do not support this hypothesis.

A second objective of this study was to estimate the timing of range expansion into the northern extents of the species' ranges. Our *F. notatus* ABC range expansion model placed coalescence between the north westerly distributed Spring River population in the Arkansas/Neosho drainage and a lower Mississippi oxbow lake population in a time period spanning the LGM (95% CI: 12,000 to 42,200 gbp), while our *F. olivaceus* ABC model placed coalescence of populations from isolated drainages in the Ozark Highlands within an overlapping time period (95% CI: 4,850-16,300). The distributions of genetic variation in both species are consistent with a coastal plain glacial refugium hypothesis and post-Pleistocene range expansion (Soltis et al., 2006).

Our sampling design was uneven across geographic regions, and insufficiently complete for a full historical reconstruction of range expansion for either species. However, some interesting inferences suggest a complex pattern of secondary contacts and admixture events during range expansions. Multiple lines of evidence supported recent admixture between the Mississippi and Red clusters in the Spring/Neosho River *F. notatus* population. The STRUCTURE and TreeMix analyses were remarkably consistent, suggesting 12 or 13% admixture from the Red cluster. However, our attempt to model this admixture event using ABC was rejected in favor of a simpler bifurcating model. We only had one *F. notatus* sample from the Arkansas River drainage system, so we are limited on what we can conclude about the history of *F. notatus* in this drainage.

Among *F. olivaceus* populations in the Red River drainage, the Glover and Kiamichi River populations formed one genetically distinct cluster, while the adjacent Cossatot River population was relatively divergent and equally closely related to *F. olivaceus* populations in several gulf coastal populations. This result points to independent colonization events of *F. olivaceus* populations in these respective tributaries that may have been facilitated by changing drainage patterns in the Ouachita Highlands (Mayden, 1985).

Phylogeographic analyses of drainages in the Ozark Highlands indicated that this region was invaded by *F. olivaceus* from two dispersal fronts. One dispersal route was along the Mississippi and Missouri River corridor, and the other route was by way of the Arkansas River and its Ozark tributaries (Fig. 7). Interestingly, Gasconade River populations exhibited admixture of both dispersal fronts, indicating that recent dispersal was facilitated by headwaters exchanges across the Ozark Plateau (Fig. 7, dotted line)(Beeson, McCoy, & Keen-Zebert, 2017). The highest levels of admixture were inferred in the lower portion of the Gasconade drainage. We modeled the Gasconade River admixture event and found strong support, with nearly equal contributions to the gene pool from each dispersal route (45% Missouri; 55% Arkansas). It is notable that St. Francis and Meramec River drainage populations cluster tightly in PCoA space, while Gasconade River drainage populations are less homogeneous. The relative homogeneity of the Meramec River drainage is consistent with a single recent colonization event, while the contrasting geographic structure in the Gasconade River supports a hypothesis of recent admixture from two invasion fronts, and ongoing migration-drift disequilibrium. Our ABC model of admixture in the Gasconade River indicates that dispersal into the Ozark Highlands occurred post-Pleistocene.

Interspecific admixture

Hybridization and introgression are recognized as important evolutionary processes that contribute to species genetic diversity and evolutionary potential (Barton & Hewitt, 1989; Dowling & DeMarais, 1993). On the other hand, evolutionary processes, like reinforcement, may drive reproductive isolation (Berdan & Fuller, 2012; Coyne & Orr, 2004). Members of the *F. notatus* complex exhibit karyotypic variation that may contribute to reduced hybrid fitness and reproductive isolation (Duvernell & Schaefer, 2014). The potential role of chromosome evolution in reproductive isolation and speciation is of interest in a wide range of species (Noor, Grams, Bertucci, & Reiland, 2001; Potter et al., 2017; Potter Sally, Moritz Craig, & Eldridge Mark D. B., 2015; Rieseberg, 2001), and remains unresolved in topminnows.

Previous studies of contact zones between *F. olivaceus* and *F. notatus* have demonstrated that hybridization can be extensive locally, but is also highly variable across geographic regions (Duvernell & Schaefer, 2014; Schaefer et al., 2016). Contact zones between topminnows isolated in separate drainages can exhibit anywhere from virtual random mating between species (e.g. $F_{IS} \approx 0$) to complete

reproductive isolation (e.g. $F_{IS} \approx 1$) (Duvernell & Schaefer, 2014). This study targeted population samples of both species outside known contact zones from eight drainages with species co-occurrence. After excluding all individuals of recent hybrid ancestry, a TreeMix analysis detected interspecific gene flow within five drainages, with admixture proportions (w) being remarkably consistent, ranging from 1.1 to 1.6%. All inferences of admixture included populations of the two species from the same drainage. Previously reported estimates of F_{IS} among drainages in which introgression was detected in this study range from 0.02 to 0.22, while the remaining drainages, in which introgression was not inferred, exhibited F_{IS} values indicative of greater reproductive isolation, between 0.41 and 0.92 (Duvernell & Schaefer, 2014). Therefore, admixture proportions estimated in this study are negatively correlated with previously estimated F_{IS} between syntopic populations ($r = -0.88$). The SNPs employed in this study have not been mapped to linkage groups, and future mapping efforts will be able to reveal if the same linkage groups have introgressed between species among isolated drainages, and the relationship of those linkage groups to chromosomal rearrangements.

Conclusions

Patterns of biodiversity are shaped by both contemporary and historical factors. Species distributions in temperate freshwater systems are determined both by environment, as well by dynamic drainage patterns and connections, which can facilitate or limit species movement (Carrea et al., 2014; Hughes, Schmidt, & Finn, 2009). Topminnows in the *Fundulus notatus* species complex are ecological generalists that are notable for their broad overlapping distributions, and limited phylogeographic structure (Duvernell & Schaefer, 2014). This study contrasted the patterns of lineage diversity for members of the *F. notatus* species complex distributed among drainages along the Gulf Coastal Plain, and throughout the Highlands of central North America. Between the two broadly distributed species, *F. notatus* occupied multiple coastal drainage refugia during the final climate cycles of the late-Pleistocene, while lack of phylogeographic structure in *F. olivaceus* suggested a single refugium. Both species underwent a northward range expansion that coincided approximately with the end of the LGM. Dispersal routes to the northern temperate portions of both species ranges would have occurred along the major branches of the contemporary Mississippi River Basin.

References

- Barton, N. H., & Hewitt, G. M. (1989). Adaptation, speciation and hybrid zones. *Nature*, 341(6242), 497–503. <https://doi.org/10.1038/341497a0>
- Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 379–406. <https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- Beeson, H. W., McCoy, S. W., & Keen-Zebert, A. (2017). Geometric disequilibrium of river basins produces long-lived transient landscapes. *Earth and Planetary Science Letters*, 475, 34–43. <https://doi.org/10.1016/j.epsl.2017.07.010>

- Bentley, S. J., Blum, M. D., Maloney, J., Pond, L., & Paulsell, R. (2016). The Mississippi River source-to-sink system: Perspectives on tectonic, climatic, and anthropogenic influences, Miocene to Anthropocene. *Earth-Science Reviews*, *153*, 139–174. <https://doi.org/10.1016/j.earscirev.2015.11.001>
- Berdan, E. L., & Fuller, R. C. (2012). A Test for Environmental Effects on Behavioral Isolation in Two Species of Killifish. *Evolution*, *66*(10), 3224–3237. <https://doi.org/10.1111/j.1558-5646.2011.01646.x>
- Black, A., & Howell, W. M. (1978). A Distinctive Chromosomal Race of the Cyprinodontid Fish, *Fundulus notatus*, from the Upper Tombigbee River System of Alabama and Mississippi. *Copeia*, *1978*(2), 280–288. <https://doi.org/10.2307/1443564>
- Blum, M. D., & Hattier-Womack, J. (2009). Climate change, sea-level change, and fluvial sediment supply to deepwater depositional systems. *SEPM Special Publication*, *92*, 15–39.
- Bossu, C. M., Beaulieu, J. M., Ceas, P. A., & Near, T. J. (2013). Explicit tests of palaeodrainage connections of southeastern North America and the historical biogeography of Orangethroat Darters (Percidae: Etheostoma: Ceasia). *Molecular Ecology*, *22*(21), 5397–5417. <https://doi.org/10.1111/mec.12485>
- Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, *26*(10), 1372–1373. <https://doi.org/10.1093/bioinformatics/btq110>
- Braasch, M. E., & Smith, P. W. (1965). Relationships of the Topminnows *Fundulus notatus* and *Fundulus olivaceus* in the Upper Mississippi River Valley. *Copeia*, *1965*(1), 46–53. <https://doi.org/10.2307/1441238>
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308>
- 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>
- Carrea, C., Anderson, L. V., Craw, D., Waters, J. M., & Burrige, C. P. (2014). The significance of past interdrainage connectivity for studies of diversity, distribution and movement of freshwater-limited taxa within a catchment. *Journal of Biogeography*, *41*(3), 536–547. <https://doi.org/10.1111/jbi.12207>
- Chen, T. R. (1971). A comparative chromosome study of twenty killifish species of the genus *Fundulus* (Teleostei: Cyprinodontidae). *Chromosoma*, *32*(4), 436–453. <https://doi.org/10.1007/BF00285254>
- 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>

- Cornuet, J.-M., Ravigné, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, *11*(1), 401. <https://doi.org/10.1186/1471-2105-11-401>
- Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland: Sinauer Associates, Inc.
- Crandall, K. A., Templeton, A. R., & Neigel, J. (1999). The Zoogeography and Centers of Origin of the Crayfish Subgenus *Procericambarus* (Decapoda: Cambaridae). *Evolution*, *53*(1), 123–134. <https://doi.org/10.1111/j.1558-5646.1999.tb05338.x>
- Cross, F. B., Mayden, R. L., & Stewart, J. D. (1986). Fishes in the western Mississippi drainage. In C. H. Hocutt & E. O. Wiley (Eds.), *Zoogeography of North American freshwater fishes* (pp. 363–412). New York: Wiley.
- Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, *25*(7), 410–418. <https://doi.org/10.1016/j.tree.2010.04.001>
- 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>
- Dowling, T. E., Broughton, R. E., & DeMarais, B. D. (1997). Significant Role for Historical Effects in the Evolution of Reproductive Isolation: Evidence from Patterns of Introgression Between the Cyprinid Fishes, *Luxilus cornutus* and *Luxilus chrysocephalus*. *Evolution*, *51*(5), 1574–1583. <https://doi.org/10.1111/j.1558-5646.1997.tb01480.x>
- Dowling, T. E., & DeMarais, B. D. (1993). Evolutionary significance of introgressive hybridization in cyprinid fishes. *Nature*, *362*(6419), 444–446. <https://doi.org/10.1038/362444a0>
- 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>
- Drury, C., Dale, K. E., Panlilio, J. M., Miller, S. V., Lirman, D., Larson, E. A., ... Oleksiak, M. F. (2016). Genomic variation among populations of threatened coral: *Acropora cervicornis*. *BMC Genomics*, *17*(1), 286. <https://doi.org/10.1186/s12864-016-2583-8>
- Duvernell, D. D., Meier, S. L., Schaefer, J. F., & Kreiser, B. R. (2013). Contrasting phylogeographic histories between broadly sympatric topminnows in the *Fundulus notatus* species complex. *Molecular Phylogenetics and Evolution*, *69*(3), 653–663. <https://doi.org/10.1016/j.ympev.2013.07.013>
- Duvernell, D. D., & Schaefer, J. F. (2014). Variation in contact zone dynamics between two species of topminnows, *Fundulus notatus* and *F. olivaceus*, across isolated drainage systems. *Evolutionary Ecology*, *28*(1), 37–53. <https://doi.org/10.1007/s10682-013-9653-z>
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. <https://doi.org/10.1007/s12686-011-9548-7>

- Echelle, A. A., Schwemm, M. R., Lang, N. J., Nagle, B. C., Simons, A. M., Unmack, P. J., ... Hoagstrom, C. W. (2014). Molecular systematics and historical biogeography of the *Nocomis biguttatus* species group (Teleostei: Cyprinidae): Nuclear and mitochondrial introgression and a cryptic Ozark species. *Molecular Phylogenetics and Evolution*, *81*, 109–119. <https://doi.org/10.1016/j.ympev.2014.09.011>
- Egge, J. J. D., & Hagbo, T. J. (2015). Comparative Phylogeography of Mississippi Embayment Fishes. *PLOS ONE*, *10*(3), e0116719. <https://doi.org/10.1371/journal.pone.0116719>
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE*, *6*(5), e19379. <https://doi.org/10.1371/journal.pone.0019379>
- Estoup, A., Lombaert, E., Marin, J.-M., Guillemaud, T., Pudlo, P., Robert, C. P., & Cornuet, J.-M. (2012). Estimation of demo-genetic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources*, *12*(5), 846–855. <https://doi.org/10.1111/j.1755-0998.2012.03153.x>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- 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>
- Gante, H. F., Micael, J., Oliva-Paterna, F. J., Doadrio, I., Dowling, T. E., & Alves, M. J. (2009). Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Molecular Ecology*, *18*(15), 3240–3255. <https://doi.org/10.1111/j.1365-294X.2009.04264.x>
- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLOS ONE*, *9*(2), e90346. <https://doi.org/10.1371/journal.pone.0090346>
- Gómez, A., & Lunt, D. H. (2007). Refugia within Refugia: Patterns of Phylogeographic Concordance in the Iberian Peninsula. In S. Weiss & N. Ferrand (Eds.), *Phylogeography of Southern European Refugia: Evolutionary perspectives on the origins and conservation of European biodiversity* (pp. 155–188). https://doi.org/10.1007/1-4020-4904-8_5
- Griffith, R. W. (1974). Environment and Salinity Tolerance in the Genus *Fundulus*. *Copeia*, *1974*(2), 319. <https://doi.org/10.2307/1442526>

- Harrington, R. C., Benavides, E., & Near, T. J. (2013). Phylogenetic Inference of Nuptial Trait Evolution in the Context of Asymmetrical Introgression in North American Darters (Teleostei). *Evolution*, 67(2), 388–402. <https://doi.org/10.1111/j.1558-5646.2012.01797.x>
- Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, 68(1–2), 87–112. <https://doi.org/10.1111/j.1095-8312.1999.tb01160.x>
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183–195. <https://doi.org/10.1098/rstb.2003.1388>
- Highton, R. (1995). Speciation in Eastern North American Salamanders of the Genus Plethodon. *Annual Review of Ecology and Systematics*, 26(1), 579–600. <https://doi.org/10.1146/annurev.es.26.110195.003051>
- Hoagstrom, C. W., Ung, V., & Taylor, K. (2014). Miocene rivers and taxon cycles clarify the comparative biogeography of North American highland fishes. *Journal of Biogeography*, 41(4), 644–658. <https://doi.org/10.1111/jbi.12244>
- Howell, W. M., & Black, A. (1981). Karyotypes in Populations of the Cyprinodontid Fishes of the *Fundulus notatus* Species-complex: A Geographic Analysis. *Bulletin of the Alabama Museum of Natural History*, (6), 19–30.
- Hughes, J. M., Schmidt, D. J., & Finn, D. S. (2009). Genes in Streams: Using DNA to Understand the Movement of Freshwater Fauna and Their Riverine Habitat. *BioScience*, 59(7), 573–583. <https://doi.org/10.1525/bio.2009.59.7.8>
- Hundt, P. J., Berendzen, P. B., & Simons, A. M. (2017). Species delimitation and phylogeography of the studfish *Fundulus catenatus* species group (Ovalentaria: Cyprinodontiformes). *Zoological Journal of the Linnean Society*, 180(2), 461–474. <https://doi.org/10.1111/zoj.12485>
- Imbrie, J., Hays, J. D., Martinson, D. G., McIntyre, A., Mix, A. C., Morley, J. J., ... Shackleton, N. J. (1984). The orbital theory of Pleistocene climate: support from a revised chronology of the marine d180 record. In A. L. Berger (Ed.), *Malinkovitch and Climate, Part 1* (pp. 269–305). Dordrecht-Holland: D. Reidel Publishing Company.
- Jeffries, D. L., Copp, G. H., Handley, L. L., Olsén, K. H., Sayer, C. D., & Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology*, 25(13), 2997–3018. <https://doi.org/10.1111/mec.13613>
- Kotlík, P., Bogutskaya, N. G., & Ekmekçi, F. G. (2004). Circum Black Sea phylogeography of *Barbus* freshwater fishes: divergence in the Pontic glacial refugium. *Molecular Ecology*, 13(1), 87–95. <https://doi.org/10.1046/j.1365-294X.2003.02021.x>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>

- Lea, D. W., Martin, P. A., Pak, D. K., & Spero, H. J. (2002). Reconstructing a 350ky history of sea level using planktonic Mg/Ca and oxygen isotope records from a Cocos Ridge core. *Quaternary Science Reviews*, *21*(1), 283–293. [https://doi.org/10.1016/S0277-3791\(01\)00081-6](https://doi.org/10.1016/S0277-3791(01)00081-6)
- Mayden, R. L. (1985). Biogeography of Ouachita Highland Fishes. *The Southwestern Naturalist*, *30*(2), 195–211. <https://doi.org/10.2307/3670734>
- Mayden, R. L. (1988). Vicariance Biogeography, Parsimony, and Evolution in North American Freshwater Fishes. *Systematic Biology*, *37*(4), 329–355. <https://doi.org/10.1093/sysbio/37.4.329>
- Mayden, R. L., & Allen, J. (2015). Phylogeography of *Pteronotropis signipinnis*, *P. euryzonus*, and the *P. hypselopterus* Complex (Teleostei: Cypriniformes), with Comments on Diversity and History of the Gulf and Atlantic Coastal Streams [Research article]. <https://doi.org/10.1155/2015/675260>
- Metcalf, A. L. (1966). *Fishes of the Kansas River system in relation to zoogeography of the Great Plains*. Lawrence: University of Kansas Museum of Natural History.
- Mittelbach, G. G., Schemske, D. W., Cornell, H. V., Allen, A. P., Brown, J. M., Bush, M. B., ... Turelli, M. (2007). Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, *10*(4), 315–331. <https://doi.org/10.1111/j.1461-0248.2007.01020.x>
- Near, T. J., Kassler, T. W., Koppelman, J. B., Dillman, C. B., & Philipp, D. P. (2003). Speciation in North American Black Basses, *Micropterus (actinopterygii: Centrarchidae)*. *Evolution*, *57*(7), 1610–1621. <https://doi.org/10.1111/j.0014-3820.2003.tb00368.x>
- Near, T. J., & Keck, B. P. (2005). Dispersal, vicariance, and timing of diversification in *Nothonotus darters*. *Molecular Ecology*, *14*(11), 3485–3496. <https://doi.org/10.1111/j.1365-294X.2005.02671.x>
- Noor, M. A. F., Grams, K. L., Bertucci, L. A., & Reiland, J. (2001). Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences*, *98*(21), 12084–12088. <https://doi.org/10.1073/pnas.221274498>
- Nunez, J. C. B., Seale, T. P., Fraser, M. A., Burton, T. L., Fortson, T. N., Hoover, D., ... Crawford, D. L. (2015). Population Genomics of the Euryhaline Teleost *Poecilia latipinna*. *PLOS ONE*, *10*(9), e0137077. <https://doi.org/10.1371/journal.pone.0137077>
- Oberdorff, T., Hugueny, B., & Guégan, J.-F. (1997). Is there an influence of historical events on contemporary fish species richness in rivers? Comparisons between Western Europe and North America. *Journal of Biogeography*, *24*(4), 461–467. <https://doi.org/10.1111/j.1365-2699.1997.00113.x>
- Pickrell, J. K., & Pritchard, J. K. (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLOS Genetics*, *8*(11), e1002967. <https://doi.org/10.1371/journal.pgen.1002967>
- Potter, S., Bragg, J. G., Blom, M. P. K., Deakin, J. E., Kirkpatrick, M., Eldridge, M. D. B., & Moritz, C. (2017). Chromosomal Speciation in the Genomics Era: Disentangling Phylogenetic Evolution of Rock-wallabies. *Frontiers in Genetics*, *8*. <https://doi.org/10.3389/fgene.2017.00010>

- Potter Sally, Moritz Craig, & Eldridge Mark D. B. (2015). Gene flow despite complex Robertsonian fusions among rock-wallaby (*Petrogale*) species. *Biology Letters*, *11*(10), 20150731. <https://doi.org/10.1098/rsbl.2015.0731>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959.
- Ray, J. M., Wood, R. M., & Simons, A. M. (2006). Phylogeography and post-glacial colonization patterns of the rainbow darter, *Etheostoma caeruleum* (Teleostei: Percidae). *Journal of Biogeography*, *33*(9), 1550–1558. <https://doi.org/10.1111/j.1365-2699.2006.01540.x>
- Rieseberg, L. H. (2001). Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, *16*(7), 351–358. [https://doi.org/10.1016/S0169-5347\(01\)02187-5](https://doi.org/10.1016/S0169-5347(01)02187-5)
- Salisbury, S. J., McCracken, G. R., Keefe, D., Perry, R., & Ruzzante, D. E. (2016). A portrait of a sucker using landscape genetics: how colonization and life history undermine the idealized dendritic metapopulation. *Molecular Ecology*, *25*(17), 4126–4145. <https://doi.org/10.1111/mec.13757>
- Schaefer, J. F., Kreiser, B. R., Champagne, C., Mickle, P. M., & Duvernell, D. D. (2009). Patterns of co-existence and hybridisation between narrowly endemic (*Fundulus euryzonus*) and broadly distributed (*F. olivaceus*) topminnows in a riverine contact zone. *Ecology of Freshwater Fish*, *18*(3), 360–368. <https://doi.org/10.1111/j.1600-0633.2009.00353.x>
- Schaefer, J. F., Duvernell, D., & Campbell, D. C. (2016). Hybridization and introgression in two ecologically dissimilar *Fundulus* hybrid zones. *Evolution*, *70*(5), 1051–1063. <https://doi.org/10.1111/evo.12920>
- Schönhuth, S., Gagne, R. B., Alda, F., Neely, D. A., Mayden, R. L., & Blum, M. J. (2018). Phylogeography of the widespread creek chub *Semotilus atromaculatus* (Cypriniformes: Leuciscidae). *Journal of Fish Biology*, *93*(5), 778–791. <https://doi.org/10.1111/jfb.13778>
- Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, *69*(1), 82–90. <https://doi.org/10.1139/z91-013>
- Snedden, J. W., Galloway, W. E., Milliken, K. T., Xu, J., Whiteaker, T., & Blum, M. D. (2018). Validation of empirical source-to-sink scaling relationships in a continental-scale system: The Gulf of Mexico basin Cenozoic record. *Geosphere*, *14*(2), 768–784. <https://doi.org/10.1130/GES01452.1>
- Soltis, D. E., Morris, A. B., McLachlan, J. S., Manos, P. S., & Soltis, P. S. (2006). Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*, *15*(14), 4261–4293. <https://doi.org/10.1111/j.1365-294X.2006.03061.x>
- Strange, R. M., & Burr, B. M. (1997). Intraspecific Phylogeography of North American Highland Fishes: A Test of the Pleistocene Vicariance Hypothesis. *Evolution*, *51*(3), 885–897. <https://doi.org/10.1111/j.1558-5646.1997.tb03670.x>
- Swift, C., Gilbert, C. R., Bortone, S. A., Burgess, G. H., & Yerger, R. W. (1986). Zoogeography of the freshwater fishes of the southeastern United States: Savannah River to Lake Pontchartrain. In C.

- H. Hocutt & E. O. Wiley (Eds.), *The Zoogeography of North American Fishes* (pp. 213–265). New York: Wiley.
- Thomerson, J. E. (1966). A comparative biosystematic study of *Fundulus notatus* and *Fundulus olivaceus* (Pisces; Cyprinodontidae). *Tulane Studies in Zoology*, *13*, 29–47.
- Vera-Escalona, I., Habit, E., & Ruzzante, D. E. (2015). Echoes of a distant time: effects of historical processes on contemporary genetic patterns in *Galaxias platei* in Patagonia. *Molecular Ecology*, *24*(16), 4112–4128. <https://doi.org/10.1111/mec.13303>
- Wallis, G. P., Cameron-Christie, S. R., Kennedy, H. L., Palmer, G., Sanders, T. R., & Winter, D. J. (2017). Interspecific hybridization causes long-term phylogenetic discordance between nuclear and mitochondrial genomes in freshwater fishes. *Molecular Ecology*, *26*(12), 3116–3127. <https://doi.org/10.1111/mec.14096>
- Wiley, E. O., & Mayden, R. L. (1985). Species and Speciation in Phylogenetic Systematics, with Examples from the North American Fish Fauna. *Annals of the Missouri Botanical Garden*, *72*(4), 596–635. <https://doi.org/10.2307/2399217>
- Yokoyama, R., Sideleva, V. G., Shedko, S. V., & Goto, A. (2008). Broad-scale phylogeography of the Palearctic freshwater fish *Cottus poecilopus* complex (Pisces: Cottidae). *Molecular Phylogenetics and Evolution*, *48*(3), 1244–1251. <https://doi.org/10.1016/j.ympev.2008.02.002>
- Zanatta, D. T., & Harris, A. T. (2013). Phylogeography and Genetic Variability of the Freshwater Mussels (Bivalvia: Unionidae) *Ellipse*, *Venustaconcha ellipsiformis* (Conrad 1836), and Bleeding Tooth, *V. Pleasii* (Marsh 1891). *American Malacological Bulletin*, *31*(2), 267–279. <https://doi.org/10.4003/006.031.0206>

Data Accessibility

All SNP data files are publicly available in Dryad database (accession numbers).