

# Riverscape community genomics: A comparative analytical approach to identify common drivers of spatial structure

Zachery D. Zbinden<sup>1</sup>  | Marlis R. Douglas<sup>1</sup> | Tyler K. Chafin<sup>1,2</sup>  | Michael E. Douglas<sup>1</sup> 

<sup>1</sup>Department of Biological Sciences, University of Arkansas, Fayetteville, Arkansas, USA

<sup>2</sup>Biomathematics and Statistics Scotland, Edinburgh, UK

## Correspondence

Zachery D. Zbinden, Department of Biological Sciences, University of Arkansas, Fayetteville, Arkansas, USA.  
Email: [zdzbnde@uark.edu](mailto:zdzbnde@uark.edu)

## Funding information

NSF Postdoctoral Research Fellowship, Grant/Award Number: DBI: 2010774

**Handling Editor:** Paul A. Hohenlohe

## Abstract

Genetic differentiation among local groups of individuals, that is, genetic  $\beta$ -diversity, is a key component of population persistence related to connectivity and isolation. However, most genetic investigations of natural populations focus on a single species, overlooking opportunities for multispecies conservation plans to benefit entire communities in an ecosystem. We present an approach to evaluate genetic  $\beta$ -diversity within and among many species and demonstrate how this riverscape community genomics approach can be applied to identify common drivers of genetic structure. Our study evaluated genetic  $\beta$ -diversity in 31 co-distributed native stream fishes sampled from 75 sites across the White River Basin (Ozarks, USA) using SNP genotyping (ddRAD). Despite variance among species in the degree of genetic divergence, general spatial patterns were identified corresponding to river network architecture. Most species ( $N = 24$ ) were partitioned into discrete subpopulations ( $K = 2$ –7). We used partial redundancy analysis to compare species-specific genetic  $\beta$ -diversity across four models of genetic structure: Isolation by distance (IBD), isolation by barrier (IBB), isolation by stream hierarchy (IBH), and isolation by environment (IBE). A significant proportion of intraspecific genetic variation was explained by IBH ( $\bar{x} = 62\%$ ), with the remaining models generally redundant. We found evidence for consistent spatial modularity in that gene flow is higher within rather than between hierarchical units (i.e., catchments, watersheds, basins), supporting the generalization of the stream hierarchy model. We discuss our conclusions regarding conservation and management and identify the 8-digit hydrologic unit (HUC) as the most relevant spatial scale for managing genetic diversity across riverine networks.

## KEYWORDS

comparative population genetics, hydrologic units, population structure, RADseq, spatial modularity, stream hierarchy model

## 1 | INTRODUCTION

Genetic diversity is a biodiversity measure that may be quantified across geography and through time (Huber et al., 2010; Leonard et al., 2017). It is tied to species' past and future evolutionary trajectories (Shelley et al., 2021). Genetic diversity is a barometer for population-level persistence in accurately reflecting demography,

connectivity, and adaptive potential (Davis et al., 2018; DeWoody et al., 2021; Paz-Vinas et al., 2018). However, genetic diversity is often underutilized in conservation planning (Laikre, 2010; Paz-Vinas et al., 2018), in part due to a suite of affiliated necessities (i.e., specialized equipment, technical expertise), all of which expand its bottom line (Blanchet et al., 2020). Moreover, when assessment does occur, it is most often limited to populations within a single species or a small

cadre of entities within a species-group, thus minimizing the potential for much-needed generalizations (Anthonysamy et al., 2018).

Therefore, when the concept of genetic diversity is applied in a comparative sense across co-distributed species, it provides a solid framework from which community-wide management and policy can be defined (Hanson et al., 2020). Systematic conservation planning (SCP; Margules & Pressey, 2000), based on managing complementary sites within a region containing unique biodiversity, could benefit from focusing on intraspecific genetic diversity measured across community members (Paz-Vinas et al., 2018; Xuereb et al., 2021). For example, multispecies assessments can reveal common dispersal barriers (Pilger et al., 2017; Roberts et al., 2013), congruent distributions of genetic diversity (Hotaling et al., 2019; Ruzich et al., 2019), relevant spatial scales for management (Blanchet et al., 2020), and associations among species characteristics and genetic diversity (Bohonak, 1999; De Kort et al., 2021; Pearson et al., 2014). Despite the potential complexity, a comprehensive and systematic management strategy can emerge, one more appropriately aligned towards managing numerous species, with long-term conservation goals beneficial to entire communities (Blanchet et al., 2017). In addition, it also tacitly encourages support by stakeholders for an overarching management plan, one representing a consensus across multiple species and ecosystems (Douglas et al., 2020).

The spatial structure—or pattern—of genetic variation within a species is primarily dictated by the interplay between gene flow and genetic drift (Holderegger et al., 2006; Hutchison & Templeton, 1999). Different observable patterns of spatial genetic structure (i.e., genetic  $\beta$ -diversity) are used to infer the influence of different underlying processes (Orsini et al., 2013). Spatial uniformity of genetic diversity (i.e., panmixia; Rosenberg et al., 2005) is the implicit null model of population structure indicative of the differentiating effects of genetic drift being overwhelmed by homogenizing gene flow. The de facto alternative is spatially continuous genetic divergence driven by an equilibrium between gene flow and drift occurring within stable, dispersal-limited populations (i.e., isolation by distance, IBD; Wright, 1943). For most species, a significant relationship between genetic dissimilarity and geographic distance is expected (Meirmans, 2012), yet the strength of this association may vary due to the intrinsic characteristics of a species (Bohonak, 1999; Singhal et al., 2018) or the extrinsic factors experienced by a species (environmental or historical) that affect dispersal (gene flow) or effective population size (genetic drift) (Orsini et al., 2013; Paz-Vinas & Blanchet, 2015). Additional models to explain spatial structure have been introduced to explore other processes. For example, genetic divergence may be further promoted by environmental dissimilarities across sites that promote local adaptation or limited/biased dispersal (isolation by environment, IBE; Bradburd et al., 2013; Wang & Bradburd, 2014). Dispersal resistance induced by physical and environmental characteristics in between sites (isolation by resistance, IBR; McRae, 2006) or barriers to dispersal (i.e., isolation by barrier, IBB; Cushman et al., 2006; Ruiz-Gonzalez et al., 2015) can also amplify the amount of genetic dissimilarity observed over a given geographic distance.

For aquatic biodiversity, patterns of genetic divergence will also be governed by the structure and architecture of the riverine network (in contemporary and past representations). Organisms within such dendritic networks are demonstrably impacted by the physical structure of the habitat, which constrains their movement (Paz-Vinas et al., 2015; Peterson et al., 2013; White et al., 2020) and leads to correspondence between genetic relatedness and the underlying structural hierarchy (Hughes et al., 2009). While this is most apparent within the contemporary structure of river networks, their historic structure, that is, paleohydrology, also serves to bookmark genetic diversity, and the effects of past connectivity or isolation can still be observed within contemporary spatial patterns (Mayden, 1988; Strange & Burr, 1997). Moreover, the hierarchical complexity of these networks will likewise dictate population processes related to colonization/extinction and effective sizes, as reflected within genetic diversities and divergences (Chiu et al., 2020; Hopken et al., 2013; Thomaz et al., 2016). Thus, spatial genetic structuring within riverine biodiversity should reflect isolation by stream hierarchy (IBH; sensu stream hierarchy model (SHM); Meffe & Vrijenhoek, 1988). The initial genesis for the SHM was narrowly defined within desert stream fishes of the American West (Meffe & Vrijenhoek, 1988). The model is an example of the more general principle of “spatial modularity” which occurs when certain sets of habitat patches are more tightly connected through individual movement than they are to others (Fortuna et al., 2009). Spatial modularity can reveal fundamental scales common among populations and possibly species, which can inform conservation strategies (Fletcher et al., 2013). Therefore, an assessment of the SHM's generality, as compared to alternative isolating regimes, was thus imperative (Brauer et al., 2018; Hopken et al., 2013).

The factors that cause genetic structure can be correlated and confounding (Meirmans, 2012; Perez et al., 2018; Wang & Bradburd, 2014). Different mechanisms can mask the occurrence of major drivers by promoting those more ancillary with regard to single-species assessments. Variation in intraspecific genetic  $\beta$ -diversity across space is driven by the balance between gene flow and genetic drift (Hutchison & Templeton, 1999), which is, in turn, tied to dispersal, life history, and biogeography (Avisé, 1992; Comte & Olden, 2018). While co-occurring species with similar histories and environments should display similar patterns of genetic structure that reflect their shared set of extrinsic factors, the degree of genetic divergence across space can vary due to differences in the species' intrinsic characteristics (Riginos et al., 2014). Weaker genetic  $\beta$ -diversity can reduce the power to model genetic structure accurately (Jones & Wang, 2012). The emerging results are two-fold: Potentially erroneous conclusions, which in turn beget ineffective management strategies. These issues can be mitigated using replicated multispecies assessments to allow influential major processes to surface, thus effectively categorizing both “signal and noise” components with the former driving patterns of regional biodiversity (Roberts et al., 2013).

Our objective was to establish an approach from which the generality of the SHM could be tested across species of a riverscape fish community. This approach would allow key drivers to be identified,

with a concurrent expectation of common processes re-emerging within these ecological networks. We accomplish this by comparing patterns of genetic diversity across 31 fish species within the White River Basin of the Ozarks (AR/MO, USA). For each, we compared four models representing major drivers of genetic structure: Isolation by distance (IBD), isolation by stream hierarchy (IBH), isolation by barrier (IBB), and isolation by environment (IBE). We predicted that IBH would consistently explain genetic  $\beta$ -diversity for most species and that variation explained by the alternative models would mostly be captured by IBH. While we expected this consistent pattern of congruence between river network architecture and genetic  $\beta$ -diversity, we expected that the degree of the strength of this association would vary among species due to their unique histories and intrinsic characteristics that influenced dispersal and effective population sizes. Our data represent thousands of SNPs (single nucleotide polymorphisms) derived via recent advances in high-throughput sequencing (Peterson et al., 2012). This technology has, in turn, allowed thousands of individuals to be genotyped as a financially and logistically practical research endeavour across multiple nonmodel species (da Fonseca et al., 2016). Our study and the data we present are relatively novel and provide an additional perspective compared to that established with meta-analytic frameworks with similar aims (e.g., Paz-Vinas et al., 2015). We offer our approach as a potential blueprint for developing more comprehensive genetic management plans at the community level.

## 2 | MATERIALS AND METHODS

### 2.1 | Study system

Our study system, the White River Basin, is located within the Western Interior Highlands of North America, a previous component of the more extensive pre-Pleistocene Central Highlands extending north and east but subsequently subsumed by numerous glacial advances into two disjunct subcomponents: Western Interior Highlands (i.e., Ozarks, Ouachita Mountains), and Eastern Highlands (i.e., Appalachian Plateau, Blue Ridge, Appalachian Highlands; Mayden, 1985). The Ozark Plateau remained an unglaciated refugium with elevated endemism and diversity (Warren et al., 2000). The White River Basin was established by at least Late Pliocene (>3 Ma; Jorgensen, 1993), but its eastern tributaries were captured by the Mississippi River when it bisected the basin during the Pleistocene (Mayden, 1988; Strange & Burr, 1997). This paleohydrologic signature may remain in contemporary patterns of population divergence in the White River Basin, as manifested by replicated patterns of genetic structure between eastern and western populations.

### 2.2 | Sampling

The sampling region for our study is composed of the White River and St. Francis River basins (AR/MO) (Figure 1). Both are tributaries

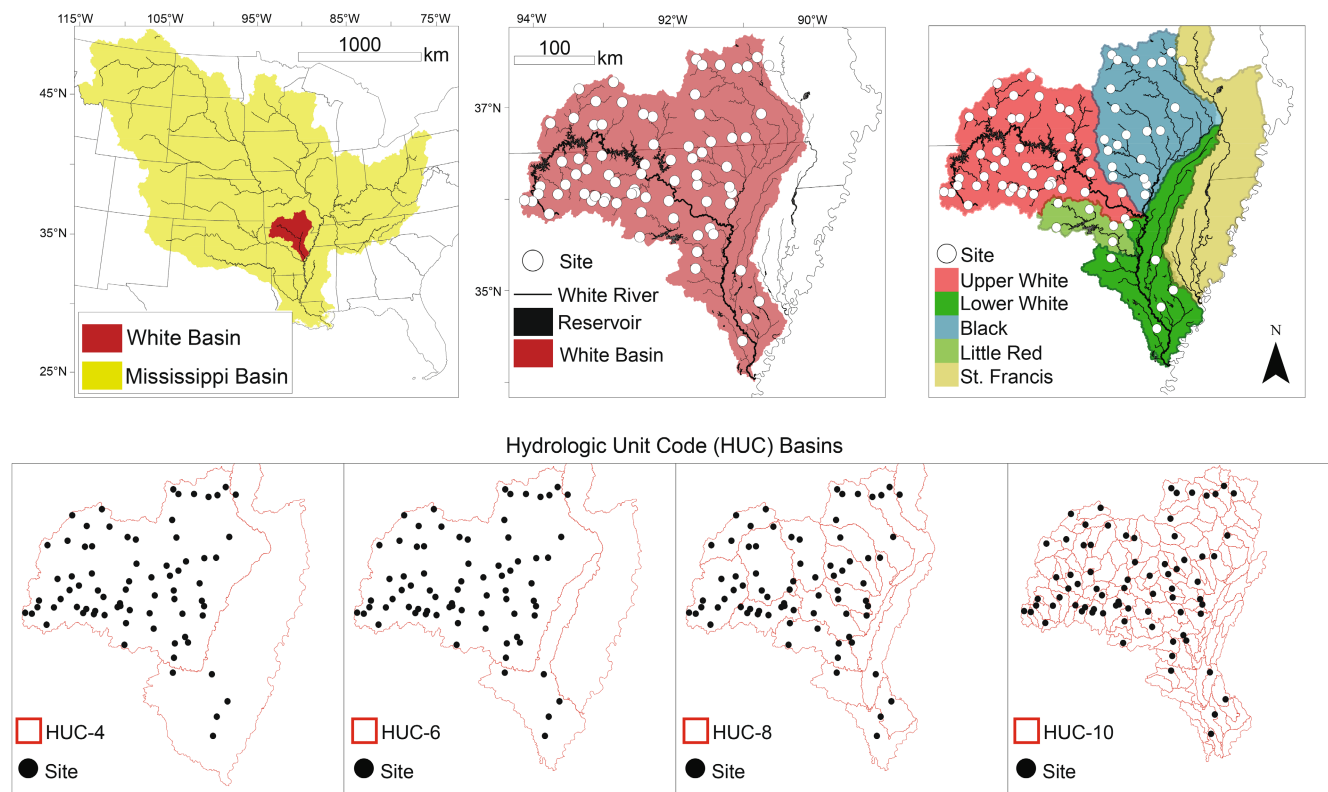
to the Mississippi River, draining 71,911 km<sup>2</sup> and 19,600 km<sup>2</sup>, respectively. Five sub-basins are apparent: St. Francis, Upper White, Black, Lower White, and Little Red rivers (Figure 1). These are further subdivided into the following hierarchical hydrologic units (HUC) (USGS & USDA-NRCS, 2013; USGS, 2021) representing different spatial scales: HUC-4 subregions ( $N = 2$ ); HUC-6 Basins ( $N = 3$ ); HUC-8 Subbasins ( $N = 19$ ); HUC-10 Watersheds ( $N = 129$ ; Figure 1).

Sampling was approved by the University of Arkansas Institutional Animal Care and Use Committee (IACUC: no. 17077), with collecting permits as follows: Arkansas Game & Fish Commission (#020120191); Missouri Department of Wildlife Conservation (no. 18136); US National Parks Service (Buffalo River Permit; BUFF-2017-SCI-0013). Fishes were sampled using seine nets in wadable streams during low flow between June 2017 and September 2018. Time spent sampling a site ranged from 30–60 min, with a target of 5–10 individuals/species encountered. Individuals were euthanized by immersion in tricaine methanesulphonate (MS-222) at a concentration of 500 mg/L, buffered to pH = 7 with subsequent preservation in 95% ethanol. Formal species diagnosis occurred in the laboratory, and the right pectoral fin was removed from each specimen and stored in 95% ethanol at  $-20^{\circ}\text{C}$  prior to subsequent DNA extraction. Specimens were housed at the Arkansas Conservation and Molecular Ecology Laboratory, University of Arkansas, Fayetteville.

### 2.3 | Genomic data collection and filtering

Genomic DNA was isolated (Qiagen Fast kits; Qiagen Inc.) and quantified by fluorometry (Qubit; Thermo-Fisher Scientific). Individuals were genotyped using double-digest restriction site-associated DNA (ddRAD) sequencing (Peterson et al., 2012), with procedures modified appropriately (Chafin et al., 2019). Standardized DNA amounts (1000 ng) were digested at  $37^{\circ}\text{C}$  with high-fidelity restriction enzymes *MspI* (5'-CCGG-3') and *PstI* (5'-CTGCAG-3') (New England Biosciences), bead-purified (Ampure XP; Beckman-Coulter Inc.), standardized to 100 ng, and then ligated with custom adapters containing in-line identifying barcodes (T4 Ligase; New England Biosciences). Samples were pooled in sets of 48 and size-selected from 326–426 bp, including adapter length (Pippin Prep; Sage Sciences). Illumina adapters and i7 index were added via 12-cycle PCR with Phusion high-fidelity DNA polymerase (New England Biosciences). Three libraries ( $3 \times 48 = 144$  individuals/lane) were pooled per lane and single-end sequenced on the Illumina HiSeq 4000 platform ( $1 \times 100$  bp; Genomics & Cell Characterization Core Facility; University of Oregon, Eugene). Quality control checks, including fragment analysis and quantitative real-time PCR, were performed at the core facility before sequencing. Raw sequence reads are deposited in the NCBI Sequence Read Archive (Zbinden et al., 2022a).

Raw Illumina reads were demultiplexed, clustered, filtered, and aligned in IPYRAD version 0.9.62 (Eaton & Overcast, 2020). Reads were first demultiplexed, allowing up to one barcode mismatch, yielding individual FASTQ files containing raw reads ( $N = 3060$  individual



**FIGURE 1** Fish were sampled at  $N = 75$  locations across the White River Basin (Ozark Mountains, USA). The study basin is contained within the larger Mississippi River Basin, and is a direct tributary to the mainstem Mississippi. The study region is subdivided into five sub-basins: Upper White, lower White, black, little red, and the St. Francis. Beyond these basins, USGS hydrologic unit codes (HUCs) were also used to characterize the stream hierarchy position of sampling locations (4-, 6-, 8-, and 10-digit HUCs) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

files). Individuals averaged  $>2$  million reads, with those extremely low removed ( $<\bar{x} - 2s$ ) to reduce errors from poor-quality sequencing. Individuals were previously screened for admixture among species using a combination of standard analyses for assigning ancestry proportions based on SNP data (Zbinden et al., 2022b). There were 70 putatively admixed individuals removed, most ( $N = 66$ ) belonging to the minnow family Leuciscidae. Raw sequence reads were partitioned by species ( $N = 31$ ) and aligned de novo in IPYRAD (Eaton & Overcast, 2020). Adapters/primers were removed, and reads with  $>5$  bases having Phred quality  $<20$  or read length  $<35$  bases (after trimming) were discarded. Clusters of homologous loci were assembled using an 85% identity threshold. Putative homologues were removed if any of the following were met:  $<20\times$  and  $>500\times$  coverage per individual;  $>5\%$  of consensus nucleotides ambiguous;  $>20\%$  of nucleotides polymorphic;  $>8$  indels present; or presence in  $<15\%$  of individuals. Paralogues were identified (and subsequently removed) as those clusters exhibiting either  $>2$  alleles per site in consensus sequence or excessive heterozygosity ( $>5\%$  of consensus bases or  $>50\%$  heterozygosity/site among individuals).

Biallelic SNP panels for each species were then visualized and filtered with the R package RADIATOR (Gosselin, 2020). To ensure high data quality, loci were removed if: Monomorphic; minor allele frequency  $<3\%$ ; mean coverage  $<20$  or  $>200$ ; missing data  $>30\%$ ; SNP position on read  $>91$ ; and if HWE was lacking in one or more

sampling sites ( $\alpha = 0.0001$ ). To reduce linkage disequilibrium, only one SNP per locus was retained (that which maximized minor allele count). Finally, singleton individuals per species at a sampling site and those with  $>75\%$  missing data in the filtered panel were removed. Species-level SNP panel alignments, metadata, and R code are available on Open Science Framework (Zbinden et al., 2022c).

## 2.4 | Exploring genetic structure

Intraspecific genetic structure was first assessed among sites for visualization, with subsequent modelling done among individuals. For each species ( $N = 31$ ), pairwise  $F_{ST}$  (Weir & Cockerham, 1984) was calculated among sites (HIERFSTAT; Goudet et al., 2017). Jost's  $D$  was also quantified among sites and globally, as it is based on the effective number of alleles rather than heterozygosity and hence less biased by sampling differences (Jost, 2008). Additional global intraspecific  $F_{ST}$  analogues were also quantified for comparison: Multiallelic  $G_{ST}$  (Nei, 1973) and unbiased  $G'_{ST}$  (Meirmans & Hedrick, 2011; MMOD; Winter, 2012). We tested for isolation by distance (IBD) using both linearized  $F_{ST}$  and Jost's  $D$ . Their relationships with river distance (log-transformed) were assessed using the Mantel test (Mantel & Valand, 1970; ECODIST; Goslee & Urban, 2020) and visualized using linear regression (Rousset, 1997).



Admixture analysis of population structure and ancestry coefficients were estimated using sparse non-negative matrix factorization (sNMF) (Frichot et al., 2014). We ran sNMF for each species, with 20 repetitions per  $K$  value (1 to  $N$  sites or 20, whichever was smallest) and  $\alpha = 100$  (LEA; Frichot & François, 2015). The best  $K$  (i.e., the number of distinct gene pools) from each sNMF run minimizes the cross-validation entropy criterion (Alexander & Lange, 2011). The best  $K$  was then used to impute missing data (impute function using method = "mode" in LEA). The sNMF algorithm was then repeated (as above) using imputed genotypes. The resulting  $Q$ -matrices of ancestry coefficients per cluster were used to map population structure and served as the isolation by barrier (IBB) model below. The indirect inference of dispersal barriers based on discrete population structure was previously used to represent IBB (Ruiz-Gonzalez et al., 2015) and is consistent with the use of population structure maps to infer the presence of dispersal barriers.

We further assessed among-site genetic variation between hydrologic units (HUCs) and discrete population clusters (determined via sNMF) using analysis of molecular variance (AMOVA) (Excoffier et al., 1992). The purpose was to determine which spatial scale of HUC explained the most variance and to compare that to the variance explained by discrete structure across  $K$  populations inferred above with sNMF. AMOVA was performed for each species at four HUC levels (4-, 6-, 8-, and 10-digit) to compare the amount of genetic variation among HUCs, all sites, and sites within HUCs. The Watershed Boundary Dataset (USGS, 2021) was used to assign HUC classifications to each site. AMOVA was then performed for each species with genetic clusters  $K > 1$  to compare the genetic variation among discrete populations, all sites, and sites within populations. Individuals with admixed ancestry between different  $K$  populations were assigned to a population based on their highest ancestry proportion. The variance components were used to estimate  $\Phi$ -statistics (analogous to  $F$ -statistics):  $\Phi_{CT}$  = the genetic variation among groups (either HUCs or discrete populations);  $\Phi_{ST}$  = the genetic variation among sites across all groups; and  $\Phi_{SC}$  = the genetic variation among sites within groups. The wrapper R package POPPR (Kamvar et al., 2015) was used to implement the PEGAS (Paradis, 2010) version of AMOVA with default settings.

## 2.5 | Modelling genetic $\beta$ -diversity

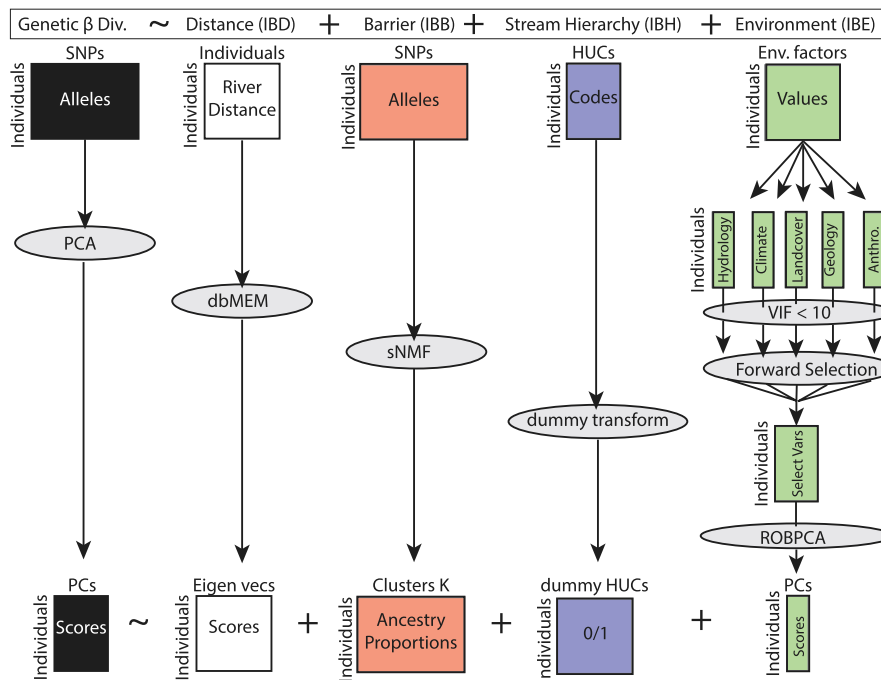
To compare four models of spatial genetic variation among groups of individuals (i.e.,  $\beta$ -diversity), we employed a variation partitioning framework based on partial redundancy analysis (Capblancq & Forester, 2021; Chan & Brown, 2020). We elected to analyse individual genetic values (i.e., individual  $\times$  SNP data matrix) rather than decomposing them into among-site distances (e.g.,  $F_{ST}$ ) to allow the use of more powerful multivariate methods rather than derived forms of the Mantel test (e.g., multiple regression on distance matrices) (Legendre & Fortin, 2010). For each species, we partitioned a matrix representing individual genetic variation among four explanatory matrices based on: IBD, IBB, IBH, and IBE (Figure 2).

Individual genetic variation within each species (i.e., individual  $\times$  SNP data matrix) was reduced to major axes of variation using principal components analysis (PCA) on each SNP panel (Xuereb et al., 2018). The appropriate number of PCs retained for each species was determined by testing the significance of observed component eigenvalues versus random eigenvalues generated by randomly shuffling the genetic data matrix and performing a PCA 999 times. The  $p$ -value for each PC axis was estimated as number of random eigenvalues equal to or larger than the observed  $+1/1000$  (Rnd-Lambda; Peres-Neto et al., 2005). This method was implemented using the R package PCDIMENSION (Coomes & Wang, 2019). For each species, individual scores on the retained PCs represented individual genetic variation (i.e., response matrix) which was modelled using the explanatory matrices (i.e., models) described below (Figure 2).

The first model (IBD) relied on river network distance measured between individuals (RIVERDIST; Tyers, 2017). The distance matrix was decomposed into positively correlated spatial eigenvectors using distance-based Moran's eigenvector maps (Borcard & Legendre, 2002; Chan & Brown, 2020; Dray et al., 2006) within the R package ADESPATIAL (Dray et al., 2020). Each individual was assigned a score for each eigenvector positively correlated with genetic diversity (Figure 2). Eigenfunction analysis is an alternative means to assess the contribution of geographic distance on patterns of genetic variation and is more powerful than Mantel correlations at detecting fine-scale structure (Xuereb et al., 2018).

The second model (IBB) was based on the discrete population structure inferred above using the admixture analysis sNMF, which was represented as the individual population coefficients (i.e.,  $Q$ -matrix). The assumption was that discrete population structure indicates a reduction of gene flow between populations due to a barrier (or high resistance) to dispersal. Similar clustering analysis output was previously used to represent IBB (Ruiz-Gonzalez et al., 2015). However, IBB often implicitly refers to instream barriers, such as dams and weirs. Our indirect approach based on population coefficients should capture the effects of those barriers as well as natural barriers to gene flow. While testing for significant genetic differences among populations defined by clustering methods is entirely meaningless due to circularity, it is reasonable to assess the strength of separation among clusters (e.g.,  $F$ -statistics from AMOVA) or variation explained by clustering (e.g., adjusted  $R^2$  from RDA) (Meirmans, 2015). Note that the IBB model could not be incorporated for species in which population structure was not apparent ( $K = 1$ ), and these species were thus tested using only three models of genetic structure.

The third model (IBH) was constructed using four levels of HUCs (4-, 6-, 8-, and 10-digit) that characterized an individual's position within the stream hierarchy, that is, hydrologic unit (USGS, 2021). Each site and individual within site could be classified by its unique HUC code at the four levels. These codes were nominal categories and were decomposed into  $N-1$  binary "dummy variables." This decomposition produces a "new" variable for each HUC within a spatial level, and an individual is either within ( $=1$ ) or outside ( $=0$ ) a given



**FIGURE 2** The analytic approach to partitioning individual genetic variation across four spatioenvironmental predictor matrices. The approach was applied separately to 31 freshwater fish species collected across the White River Basin of the Ozarks, USA. The full redundancy analysis model is shown at the top of the figure, where genetic diversity is explained by geographic distance (IBD), discrete population structure (IBB), stream hierarchical position (IBH), and environmental variation among habitats (IBE). The initial data matrices representing genetic  $\beta$ -diversity (i.e., response variable) and the four explanatory variables sets are depicted at the top. Each matrix is labelled to show rows, columns, and values (e.g., individuals, single nucleotide polymorphisms, and alleles). These matrices each pass through analyses and/or transformations (grey ellipses) to yield the matrices used for modelling at the bottom of the figure. dbMEM, distance-based Moran's eigenvector maps; dummy transform, transforming categorical variable into separate binary variables; HUCs, hydrologic units; PCA, principal component analysis; PCs, principal components; sNMF, sparse non-negative matrix factorization; ROBPCA, robust principal components analysis; SNP, single nucleotide polymorphism; VIF, variance inflation factor. Note that environmental factors were standardized (z-score) [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/mec.16806)]

HUC. Note: the dummy variables produced were one less than the number of original categories to avoid introducing collinearity.

The fourth model (IBE) relied on contrasting environmental variation across sites—represented using principal component scores (Figure 2). Environmental variables were taken from a compendium of 281 factors distinguished by five major classes: (i) Hydrology/physiography, (ii) climate, (iii) land cover, (iv) geology, and (v) anthropogenic (HYDRORIVERS version 1.0; Linke et al., 2019). We aimed to reduce the environmental factors to include only those demonstrating a significant relationship with genetic variation. Variables were first analysed within major classes, with invariant factors and those exhibiting collinearity being removed in a stepwise manner (USDM; Naimi, 2013) until each had a variation inflation factor (VIF) <10 to reduce collinearity. Standardization occurred by subtracting means and dividing by standard deviations to ensure variables had a common scale. Variables within each major class (e.g., climate) were selected for subsequent analyses using forward selection (Blanchet et al., 2008). Variables within the class being considered (e.g., climate) were first tested for a significant relationship with the response data (individual genetic variation) using redundancy analysis (RDA; Rao, 1964). If the overall relationship was significant ( $\alpha < 0.05$ ), a stepwise forward procedure

was carried out such that specific variables were selected if the adjusted  $R^2$  of the model increased significantly ( $\alpha < 0.05$ ) and the adjusted  $R^2$  did not exceed that of the overall model (Blanchet et al., 2008). This procedure was employed using the `ordiR2step` function in the R package `VEGAN` (Oksanen et al., 2020). The selected variables from each of the five classes (hydrology/physiography, climate, land cover, geology, and anthropogenic) were combined into a single environmental matrix, then reduced to a set of PCs using robust principal components analysis (ROBPCA; Hubert et al., 2005). The number of PCs retained was determined following Hubert et al. (2005), as implemented in the R package `ROSPCA` (Hubert et al., 2016). The environmental matrix of individuals  $\times$  PCs was used to model IBE.

For each species, individual genetic variation (individuals  $\times$  PCs matrix) was partitioned among the four explanatory models of genetic structure (IBD, IBB, IBH, IBE) using variation partitioning based on partial redundancy analysis (pRDA; Anderson & Legendre, 1999; Borcard et al., 1992). Redundancy analysis (Rao, 1964) is a constrained ordination technique and an extension of multiple regression that summarizes the relationship between linear combinations of multiple response variables and linear combinations of multiple explanatory variables (Capblancq & Forester, 2021). The overall

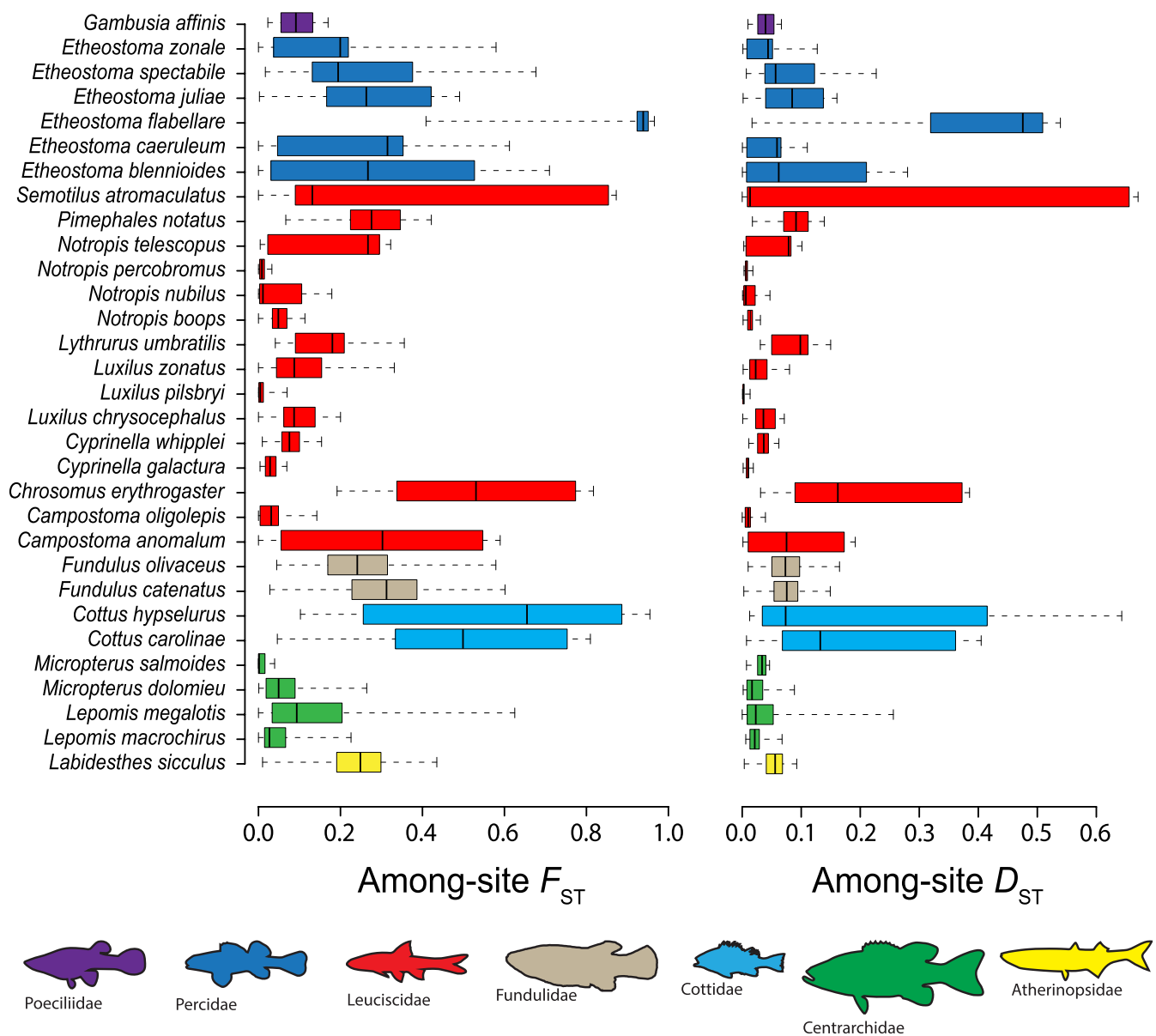
**TABLE 1** Fish species ( $N = 31$ ) were collected at 75 sampling locations across the White River Basin of the Ozark Mountains, USA. Summary data are tabulated for  $N = 2861$  individuals across seven families genotyped and analysed for genetic structure

Family	Species	NI	NS	NI/S	Reads	Loci	Depth	Ho	SNPs	Miss	PCs	PCvar
Atherinopsidae	<i>Labidesthes sicculus</i>	99	18	5.5	2,401,513	19,532	83	0.0013	2956	0.11	17	40.2
Centrarchidae	<i>Lepomis macrochirus</i>	63	17	3.7	2,369,445	26,142	61	0.0028	5873	0.14	19	45.5
	<i>Lepomis megalotis</i>	242	44	5.5	2,330,434	25,126	59	0.0036	4841	0.13	48	45.2
	<i>Micropterus dolomieu</i>	56	15	3.7	2,014,858	21,420	58	0.0018	2813	0.11	11	32.6
	<i>Micropterus salmoides</i>	15	7	2.1	2,338,155	22,827	65	0.0018	2825	0.06	7	59.4
Cottidae	<i>Cottus caroliniae</i>	24	9	2.7	2,973,760	27,523	74	0.0012	5798	0.12	5	61.6
	<i>Cottus hypselurus</i>	40	8	5.0	3,226,846	28,108	76	0.0015	7116	0.11	5	75.1
Fundulidae	<i>Fundulus catenatus</i>	112	23	4.9	2,757,508	30,509	52	0.0014	3378	0.13	18	46.0
	<i>Fundulus olivaceus</i>	131	24	5.5	2,647,685	27,631	51	0.0025	3111	0.14	22	42.6
Leuciscidae	<i>Campostoma anomalum</i>	93	20	4.7	2,226,556	16,753	77	0.0036	3187	0.13	10	36.7
	<i>Campostoma oligolepis</i>	119	31	3.8	2,038,589	16,107	76	0.0030	3121	0.12	40	44.7
	<i>Chrosomus erythrogaster</i>	53	7	7.6	2,180,045	16,508	73	0.0033	3440	0.14	6	55.8
	<i>Cyprinella galactura</i>	72	10	7.2	1,648,530	14,839	72	0.0029	3322	0.11	27	52.1
	<i>Cyprinella whipplei</i>	29	6	4.8	1,870,427	14,599	84	0.0033	2847	0.12	8	39.5
	<i>Luxilus chrysocephalus</i>	57	13	4.4	1,677,176	15,089	68	0.0025	2168	0.14	17	47.2
	<i>Luxilus pilsbryi</i>	244	31	7.9	2,028,625	16,063	81	0.0033	4922	0.14	93	52.1
	<i>Luxilus zonatus</i>	98	16	6.1	2,273,167	16,964	89	0.0030	5496	0.12	12	24.7
	<i>Lythrurus umbratilis</i>	23	5	4.6	1,970,516	16,465	68	0.0032	2491	0.12	6	40.3
	<i>Notropis boops</i>	233	28	8.3	2,355,581	15,684	104	0.0040	6161	0.11	71	43.8
	<i>Notropis nubilis</i>	191	32	6.0	2,087,695	15,544	81	0.0040	4018	0.14	65	46.3
	<i>Notropis percobromus</i>	62	10	6.2	2,082,050	17,852	74	0.0047	4393	0.13	36	65.6
	<i>Notropis telescopus</i>	81	13	6.2	2,092,015	16,154	85	0.0024	4741	0.11	12	31.2
	<i>Pimephales notatus</i>	47	13	3.6	2,106,907	15,271	92	0.0029	4022	0.13	11	49.3
	<i>Semotilus atromaculatus</i>	30	9	3.3	2,216,336	15,406	84	0.0020	2644	0.15	2	63.6
Percidae	<i>Etheostoma biennioides</i>	52	14	3.7	2,491,915	21,416	71	0.0024	5124	0.11	2	36.4
	<i>Etheostoma caeruleum</i>	358	50	7.2	2,170,268	21,900	62	0.0044	3511	0.13	20	28.7
	<i>Etheostoma flabellare</i>	22	6	3.7	2,288,120	21,041	62	0.0015	9927	0.08	4	88.7
	<i>Etheostoma juliae</i>	57	10	5.7	2,513,876	20,652	84	0.0014	5473	0.1	7	39.5
	<i>Etheostoma spectabile</i>	49	10	4.9	2,565,769	23,873	64	0.0051	5519	0.15	6	33.6
	<i>Etheostoma zonale</i>	74	15	4.9	2,364,158	21,514	74	0.0033	10,033	0.13	5	24.9
Poeciliidae	<i>Gambusia affinis</i>	35	8	4.4	2,657,603	24,021	78	0.0021	3818	0.09	9	39.9
	MEAN	92.3	16.8	5.1	2,289,230.0	20,081.7	73.6	0.0028	4486.7	0.12	20.0	46.2
	STDEV	80.8	11.2	1.5	341,159.5	4697.6	12.0	0.0010	1931.1	0.02	22.1	14.3

Abbreviations: Depth, mean coverage of loci; Family, fish family;  $H_o$ , mean observed heterozygosity; Loci, mean number of loci recovered by iPyRAD; Miss, mean missing data; NI, number of individuals analysed after filtering; NS, number of sites at which filtered individuals occurred; NI/S, mean number of individuals per site; PCs, number of principal components used to characterize neutral genetic variation; PCvar, the original genetic variance explained by the set of PCs; Reads, mean number of raw reads recovered by Illumina HiSeq; SNPs, number of single nucleotide polymorphisms in the analysed data panel; Species, species name.

variance in the response matrix is partitioned into constrained and unconstrained fractions. The former is interpreted as the amount of response variation “explained” by the explanatory set (s), expressed as a proportion equivalent to  $R^2$  in multiple regression, which is adjusted to reduce bias introduced by multiple predictors (Peres-Neto et al., 2006). The formula for the full RDA model herein was: Genetic  $\beta$ -diversity ~ IBD + IBB + IBH + IBE (Figure 2). Variation partitioning consists of running sequential pRDAs to adjust the linear effects of an explanatory set on the response by accounting for other explanatory sets first, for example, genetic  $\beta$ -diversity ~ IBD | (IBB + IBH + IBE). In other words, here we estimate variation explained by distance after removing that explained by the other three response matrices. This set of tests allows partitioning the variation explained into that “purely” attributable to a given explanatory set

(e.g., IBD) and into shared components attributable to two or more sets (e.g., IBD + IBH). Variation partitioning allows the correlation structure among competing models to be visualized, typically using Venn diagrams showing adjusted  $R^2$  for different intersectionalities of the explanatory data sets (i.e., pure and correlated components). The variation partitioning was performed using the varpart function in the R package VEGAN (Oksanen et al., 2020). The models were judged based on their significance, adjusted coefficients of determination  $R^2$ , and correlation with other models. The “best” model should explain the most variation in the genetic response matrix (i.e., highest adj.  $R^2$ ), should account for variation after accounting for that explained by other models (i.e., “pure” adj.  $R^2 > 0$ ), and account for most variation explained by competing models (i.e., other models are redundant).



**FIGURE 3** Genetic structure of  $N = 31$  fish species collected across the White River basin (Ozark Mountains, USA) as summarized by among-site  $F_{ST}$  (Weir and Cockerham's  $\theta$ ) and Jost's  $D$ . Boxplots show the distributions of both pairwise estimates among sampling sites for each species. Inner quantiles are coloured to indicate species in the same family ( $N = 7$ ) [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/mec.16806)]

### 3 | RESULTS

#### 3.1 | Sampling and data recovery summarized

Collections ( $N = 75$ ; Figure 1) yielded  $N = 72$  species and  $N = 3605$  individuals. On average, we collected  $\sim 11$  species/site, typical for streams sampled with seine nets in North America (Matthews, 1998) and similar highland streams within the Mississippi Basin (Zbinden, Geheber, Lehrter, & Matthews, 2022; Zbinden, Geheber, Matthews, & Marsh-Matthews, 2022).

We genotyped  $N = 3060$  individuals across  $N = 31$  species, with at least two individuals collected at  $\geq 5$  sampled sites. Simulations and empirical evaluations underscore the accuracy of  $F_{ST}$  estimates when large numbers of SNPs ( $\geq 1500$ ) are employed across a minimum of two individuals (Nazareno et al., 2017; Willing et al., 2012). After removing samples with missing data  $>75\%$  and those as singletons of their species at a site, the remaining  $N = 2861$  were analysed for genetic structure (Table 1). The number of individuals analysed per species ranged from 15–358 ( $\bar{x} = 92.3$ ;  $s = 80.8$ ), and the sites at which each species was collected ranged from 5–50 ( $\bar{x} = 16.8$ ;

TABLE 2 Summary of genetic structure observed for  $N = 31$  species of fish collected across the White River Basin, USA. Classifications to family and species are provided for each, along with summaries of genetic structure

Family	Species	$H_T$	$H_S$	$G_{ST}$	$G'_{ST}$	$D$	IBD	Structure	Model	Model Var
Percidae	<i>Etheostoma flabellare</i>	0.35	0.02	0.93	0.96	0.40	–	X	Stream hierarchy	99%
Leuciscidae	<i>Semotilus atromaculatus</i>	0.30	0.09	0.70	0.79	0.26	X	X	Stream hierarchy	91%
Cottidae	<i>Cottus hypselurus</i>	0.24	0.07	0.73	0.81	0.22	–	X	Stream hierarchy	99%
Leuciscidae	<i>Chrosomus erythrogaster</i>	0.27	0.11	0.59	0.71	0.21	X	X	Stream hierarchy	98%
Cottidae	<i>Cottus carolinae</i>	0.26	0.11	0.58	0.69	0.19	X	X	Stream hierarchy	93%
Leuciscidae	<i>Campostoma anomalum</i>	0.20	0.12	0.38	0.45	0.09	X	X	Stream hierarchy	87%
Percidae	<i>Etheostoma blennioides</i>	0.21	0.13	0.35	0.43	0.09	X	X	Stream hierarchy	98%
Leuciscidae	<i>Pimephales notatus</i>	0.25	0.18	0.28	0.36	0.09	X	X	Stream hierarchy	98%
Percidae	<i>Etheostoma juliae</i>	0.23	0.16	0.29	0.37	0.09	X	X	Stream hierarchy	97%
Leuciscidae	<i>Lythrurus umbratilis</i>	0.30	0.25	0.17	0.27	0.09	–	–	Stream hierarchy	69%
Percidae	<i>Etheostoma spectabile</i>	0.20	0.14	0.31	0.38	0.08	X	X	Stream hierarchy	99%
Fundulidae	<i>Fundulus olivaceus</i>	0.24	0.18	0.25	0.32	0.08	X	X	Stream hierarchy	88%
Fundulidae	<i>Fundulus catenatus</i>	0.20	0.14	0.31	0.37	0.07	X	X	Stream hierarchy	83%
Atherinopsidae	<i>Labidesthes sicculus</i>	0.18	0.14	0.24	0.29	0.05	X	X	Stream hierarchy	84%
Leuciscidae	<i>Notropis telescopus</i>	0.20	0.16	0.20	0.25	0.05	X	X	Stream hierarchy	60%
Percidae	<i>Etheostoma caeruleum</i>	0.14	0.10	0.27	0.30	0.04	X	X	Stream hierarchy	90%
Percidae	<i>Etheostoma zonale</i>	0.16	0.13	0.20	0.25	0.04	X	X	Stream hierarchy	98%
Leuciscidae	<i>Luxilus chrysocephalus</i>	0.26	0.23	0.11	0.15	0.04	X	X	Stream hierarchy	38%
Centrarchidae	<i>Lepomis megalotis</i>	0.18	0.15	0.17	0.21	0.04	X	X	Stream hierarchy	47%
Poeciliidae	<i>Gambusia affinis</i>	0.26	0.24	0.10	0.14	0.04	X	X	Stream hierarchy	59%
Leuciscidae	<i>Cyprinella whipplei</i>	0.26	0.24	0.09	0.14	0.04	X	X	Stream hierarchy	50%
Centrarchidae	<i>Micropterus salmoides</i>	0.30	0.28	0.06	0.10	0.03	X	–	Stream hierarchy	12%
Leuciscidae	<i>Luxilus zonatus</i>	0.19	0.17	0.11	0.14	0.03	–	X	Stream hierarchy	76%
Centrarchidae	<i>Lepomis macrochirus</i>	0.24	0.22	0.07	0.10	0.02	–	–	Stream hierarchy	19%
Centrarchidae	<i>Micropterus dolomieu</i>	0.23	0.22	0.07	0.10	0.02	X	–	Stream hierarchy	57%
Leuciscidae	<i>Notropis boops</i>	0.17	0.16	0.06	0.08	0.01	X	X	Stream hierarchy	23%
Leuciscidae	<i>Notropis nubilus</i>	0.14	0.13	0.07	0.08	0.01	X	X	Stream hierarchy	13%
Leuciscidae	<i>Campostoma oligolepis</i>	0.17	0.16	0.05	0.06	0.01	X	X	Stream hierarchy	15%
Leuciscidae	<i>Cyprinella galactura</i>	0.18	0.18	0.04	0.05	0.01	–	–	Stream hierarchy	12%
Leuciscidae	<i>Notropis percobromus</i>	0.18	0.18	0.03	0.04	0.01	X	–	Stream hierarchy	3%
Leuciscidae	<i>Luxilus pilsbryi</i>	0.14	0.13	0.02	0.02	0.00	X	–	Stream hierarchy	6%
	MEAN	0.22	0.16	0.25	0.30	0.08				63%
	STDEV	0.05	0.06	0.23	0.25	0.09				35%

Note: Species are ordered by Jost's  $D$ .

Abbreviations:  $D$ , Jost's genetic differentiation;  $G_{ST}$ , Nei's fixation index;  $G'_{ST}$ , unbiased fixation index;  $H_S$ , within-site heterozygosity;  $H_T$ , total heterozygosity; IBD, significant tests of isolation by distance denoted "X"; Model Var, the amount of variance explained by the best isolation model; Structure, whether the species could be subdivided into more than one population, denoted "X"; Model, the isolation model explaining the most individual genetic variance.



$s = 11.2$ ). The number of individuals/species/site ranged from 2–15 ( $\bar{x} = 5.1$ ;  $s = 1.5$ ). The mean number of raw reads/individual/species spanned from 1.65 million to 3.22 million ( $\bar{x} = 2,289,230.0$ ;  $s = 341,159.5$ ). The mean  $N$  of loci/species recovered by IPYRAD ranged from 14,599–30,509 ( $\bar{x} = 20,081.7$ ;  $s = 4697.6$ ) with a mean sequencing depth/locus of  $73.6\times$  ( $s = 12.0\times$ ). After filtering loci and retaining one SNP per locus, the panels for each species contained 2168–10,033 polymorphic sites ( $\bar{x} = 4486.7$ ;  $s = 1931.1$ ) with mean missing data/species at 12% ( $s = 2\%$ ).

### 3.2 | Genetic structure

#### 3.2.1 | Among-site genetic divergence

Distributions of among-site  $F_{ST}$  and  $D$  varied widely among species (Figure 3; Table S1), as did global indices of genetic divergence (Table 2). All three global indices of fixation or genetic divergence ( $G_{ST}$ ,  $G''_{ST}$ ,  $D$ ) were negatively correlated with within-site heterozygosity ( $H_S$ ), positively correlated with total heterozygosity ( $H_T$ ), and highly, positively correlated with each other (Table 3).

A significant relationship was found between linearized among-site  $F_{ST}$  and log-transformed among-site river network distance for 23 (74%) of the  $N = 31$  species (Figure 4). Mantel coefficients ranged from 0.11–0.88 ( $\bar{x} = 0.51$ ;  $s = 0.19$ ). Results were largely similar when IBD was tested with Jost's  $D$ , again with the same 23 species showing a significant relationship, along with two additional taxa: Smallmouth Bass (*Micropterus dolomieu*; Lacepède, 1802) and Largemouth Bass (*Micropterus salmoides*; Lacepède, 1802). Mantel correlation coefficients ranged from 0.15–0.92 ( $\bar{x} = 0.51$ ;  $s = 0.19$ ).

#### 3.2.2 | Population structure

An apparent lack of discrete genetic structure emerged across seven species, suggesting continuous structuring at the spatial scale of our study (Figure 5). For the remaining 24 species, at least two and up to seven discrete subpopulations were identified (Figure 6). This structure corresponded at the broadest hierarchical level to the two major northern basins: Upper White and Black rivers, for all species sampled in both sub-basins ( $N = 22$ ). There was also evidence of fine-scale structure for five species within the Little Red River Basin. Smaller catchments with distinct gene pools across multiple species included: North Fork (4 spp.), Buffalo (3 spp.), Upper Black (4 spp.), Current (3 spp.), and Spring rivers (4 spp.).

#### 3.2.3 | AMOVA

Discrete genetic structuring was also supported via AMOVA. Genetic variation among HUCs was significant for 24 species (Table 4). For the other seven species, variation among HUCs was  $\leq 1\%$ , save for Ozark Sculpin (*Cottus hypselurus*; Robins & Robison, 1985) and Creek

Chub (*Semotilus atromaculatus*; Mitchill, 1818). HUC differences for these accounted for  $>80\%$  of the genetic variance but were nonsignificant due to a lack of power. Southern Redbelly Dace (*Chrosomus erythrogaster*; Rafinesque, 1820) could not be tested due to a lack of repeated samples within HUC levels. Further evidence of genetic structure among HUCs was revealed in the pattern of  $\Phi_{SC}$  (genetic divergence among sites within HUCs)  $< \Phi_{ST}$  (divergence among all sites) found across 26 species. The 8-digit HUC level explained the greatest genetic variance across 21 species (Table 4).

Genetic variation among discrete population clusters (based on sNMF) was significant for 21 of the  $N = 31$  species (Table 4). Seven species were best described as single populations ( $K = 1$ ) and were not tested. The three species without significant structure, despite  $K > 1$  via sNMF, could probably be explained by low power resulting from a small number of sample sites. Again, as with HUCs,  $\Phi_{SC} < \Phi_{ST}$  was observed. However, all tested species showed this pattern (i.e., sites within the same population were less differentiated than sites across all populations).

### 3.3 | Modelling genetic $\beta$ -diversity

Variability in genetic  $\beta$ -diversity was partitioned across four models of genetic structure for the  $N = 31$  species. Principal components of SNP panel variation served as representatives of genetic variation. Across species, the number of genetic PCs ranged from 2–93 ( $\bar{x} = 20.0$ ;  $s = 20.1$ ; Table 1).

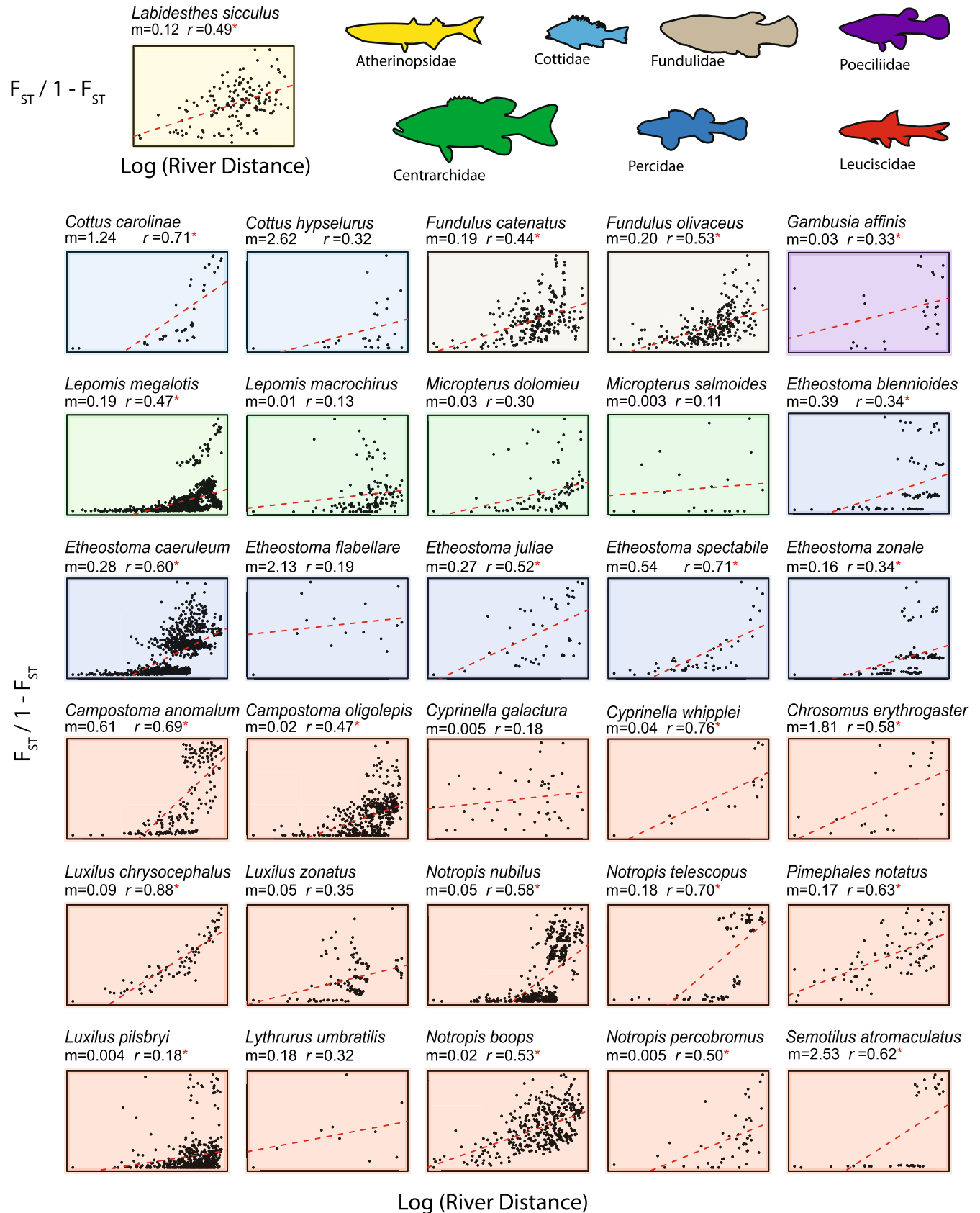
Combining the four models (IBD, IBB, IBH, IBE) explained between 3% and 100% of the genetic  $\beta$ -diversity across species ( $\bar{x} = 63.0\%$ ;  $s = 35.3\%$ ; Figure 7). Isolation by stream hierarchy (IBH;  $\bar{x} = 62.0\%$ ;  $s = 34.7\%$ ) and barrier (IBB;  $\bar{x} = 49.3\%$ ;  $s = 30.0\%$ ) contributed most to the total variation explained, while distance (IBD;  $\bar{x} = 32.1\%$ ;  $s = 25.1\%$ ) and environment (IBE;  $\bar{x} = 33.0\%$ ;  $s = 21.4\%$ ) explained less (Figure 7; Table S1). Variation explained by “pure” models, after accounting for that explained by the other three, was  $>0$  only for stream hierarchy and barrier (Figure 7; Table S1), suggesting that distance and environment are encapsulated by the former. Indeed, correlative structure among

TABLE 3 Summary of correlation among population genetic parameter estimates calculated for  $N = 31$  fish species collected across the White River Basin, USA

	$H_S$	$H_T$	$G_{ST}$	$G''_{ST}$
$H_T$	ns	–		
$G_{ST}$	–0.75	0.52	–	
$G''_{ST}$	–0.71	0.55	0.99	–
$D$	–0.65	0.67	0.97	0.96

Note: Pearson's product-moment correlation between each parameter estimate is shown in the table. Only significant ( $\alpha < 0.05$ ) correlations are shown.

Abbreviations:  $D$ , Jost's genetic differentiation;  $G_{ST}$ , Nei's fixation index;  $G''_{ST}$ , unbiased fixation index;  $H_T$ , total heterozygosity;  $H_S$ , within-site heterozygosity.



**FIGURE 4** Isolation by distance plots for  $N = 31$  fish species collected across the White River Basin (Ozark Mountains, USA). Each depicts the relationship between among-site  $F_{ST}$  (linearized) and log river distance among sites. The following are represented below each species name:  $m$  = slope of the linear regression model (dashed red line) and  $r$  = the Mantel coefficient indicating the strength of the correlation between genetic structure and distance. Significant  $r$ -values denoted with a red asterisk ( $\alpha \leq 0.05$ ) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

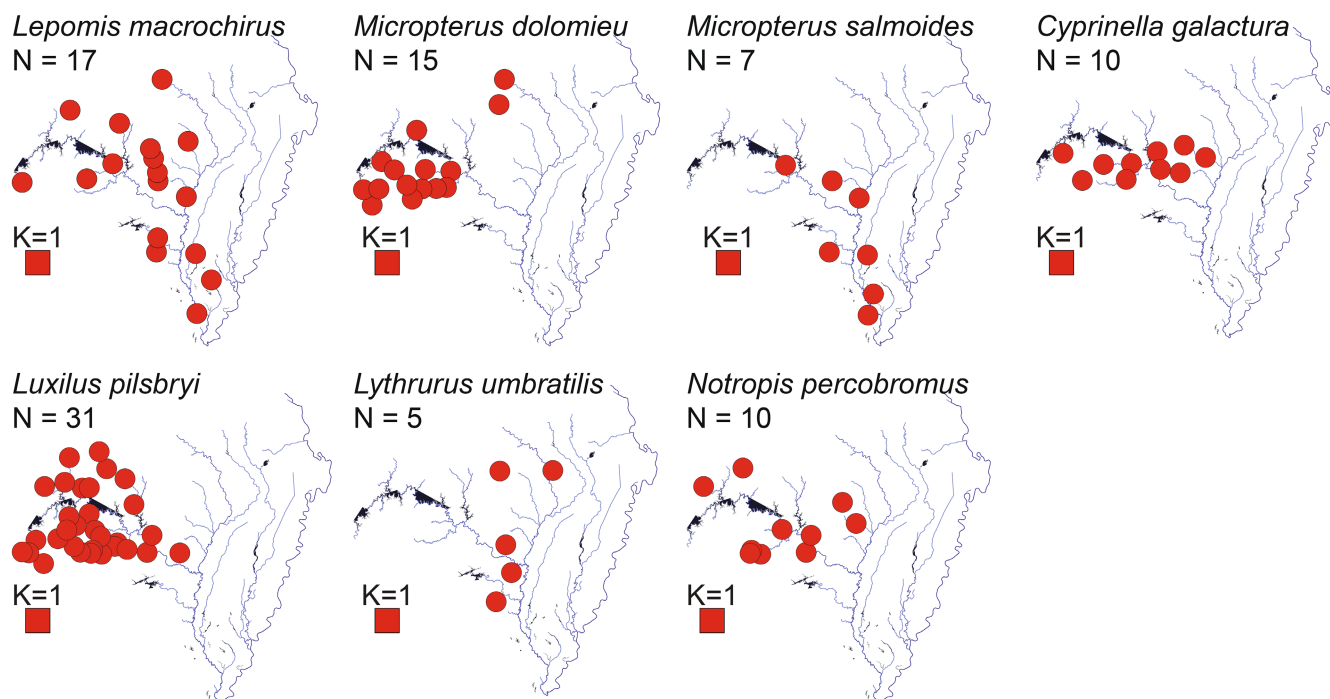
models revealed most genetic variance was explained by stream hierarchy, with the other models largely redundant (Figure 8; Table S1).

#### 4 | DISCUSSION

Genetic diversity is an essential metric for inferring evolutionary processes and guiding conservation. Single-species estimates of genetic diversity are standard given practical constraints, for example, funding mandates for species of conservation concern. However, adopting a multispecies approach for analysing genetic diversity could allow for more comprehensive and systematic management plans to be developed by focusing on commonalities (rather than differences) among species. The stream hierarchy model (Meffe & Vrijenhoek, 1988) posits that the dispersal of stream-dwelling organisms is more limited between hierarchical units (basins, sub-basins, watersheds) than within (i.e., spatial modularity; Fortuna et al., 2009). If this model was generalizable, it could determine relevant scales and regions for managing genetic diversity that may harbour complementary biodiversity and aid in systematic conservation planning (Margules & Pressey, 2000; Paz-Vinas et al., 2018; Xuereb et al., 2021).

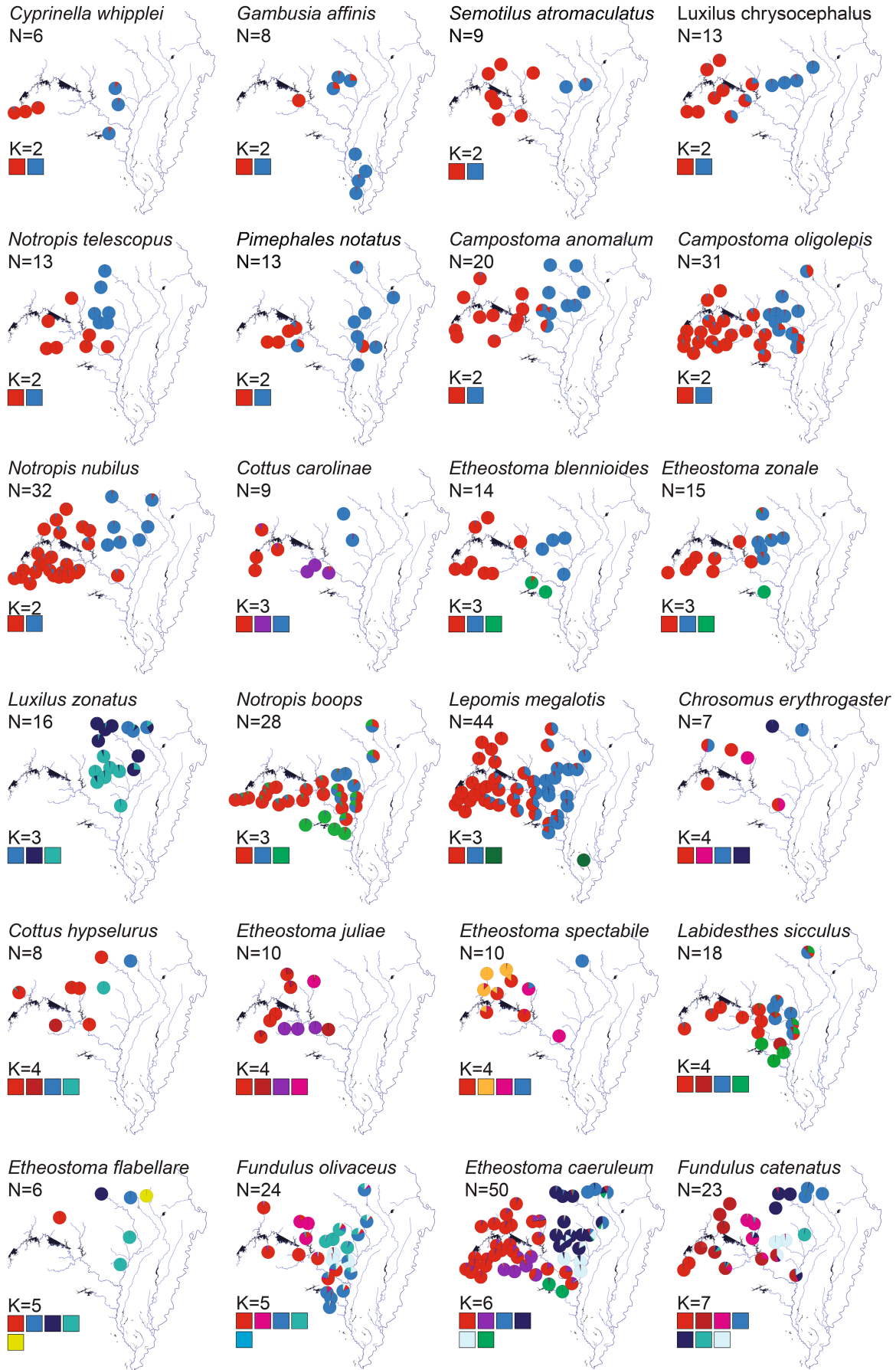
Few studies have analytically compared the spatial structure of genetic  $\beta$ -diversity. Paz-Vinas et al. (2018) compared genetic  $\alpha$ - and  $\beta$ -diversity among six river-dwelling fish species in France using microsatellites. In contrast to our results, their study found that species did not conform to common spatial patterns, although their approach differed from that herein (i.e., "hot"- and "coldspots" of  $\alpha$ - and  $\beta$ -diversity). Fortuna et al. (2009) used isozymes to demonstrate that three of four Mediterranean shrubs in their study displayed similar patterns of network connectivity (modularity) in that the same sets of sampling sites formed consistent patches based on estimates of genetic connectivity (i.e., higher connectivity among sites within a patch) that probably represent fundamental scales for populations.

Our multispecies approach yielded two salient points: (1) From a macro-perspective, river network topology and complexity are manifested in common patterns of genetic structure across species (consistent modularity); and (2) on a finer scale, the degree of intraspecific genetic divergence varies widely among codistributed species. Most species showed significant IBD patterns but also discrete population substructure, as reflected most strongly by sub-basin delineations (e.g., HUC-8). These patterns were corroborated by AMOVA and variation partitioning and are generalized across



**FIGURE 5** Sampling distribution maps of seven species which showed no evidence of discrete genetic population structure within the White River Basin (Ozark Mountains, USA). A total of  $N = 31$  species were sampled across 75 sites. The number of collection sites (red circles) for each species is denoted by  $N$ ;  $K = 1$  is the number of discrete genetic populations discerned from sparse non-negative matrix factorization [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/mec.16806)]

**FIGURE 6** Sampling distribution maps of 24 species which showed evidence of genetic population structure within the White River Basin (Ozark Mountains, USA).  $N = 31$  species were sampled across 75 sites.  $K =$  the number of discrete genetic populations discerned from sparse non-negative matrix factorization. Sampling sites are denoted as pie charts representing the average population coefficients for each site.  $N =$  number of sites where each species was collected [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/mec.16806)]

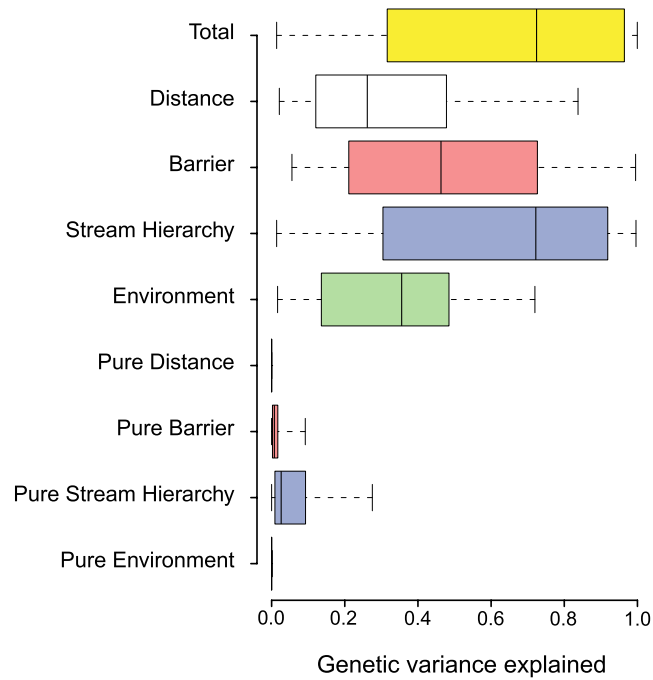


**TABLE 4** Genetic variation of fish species (N = 31) sampled across the White River Basin (Ozark Mountains, USA), was tested using analysis of molecular variance (AMOVA) to determine the proportion of genetic variation differing among distinct hydrologic units (HUCs) and among discrete population clusters. HUC tests were performed at four HUC-levels (4-, 6-, 8-, and 10-digit HUCs) and the level depicting the most genetic variance is shown

Family	Species	Hydrologic units										Population clusters					
		HUC-level	Among HUCs			Among sites			Among pops			Among sites					
			%var	sig.	$\Phi_{ST}$	%var	$\Phi_{ST}$	sig.	$\Phi_{SC}$	%var	sig.	$\Phi_{ST}$	%var	sig.	$\Phi_{SC}$		
Atherinopsidae	<i>Labidesthes sicculus</i>	HUC-8	21%	*	0.40	0.24	0.40	25%	*	0.436	18%	*	0.243				
Centrarchidae	<i>Lepomis macrochirus</i>	-	0%	ns	0.07	0.07	-	-	-	-	-	-	-				
	<i>Lepomis megalotis</i>	HUC-4	70%	*	0.77	0.23	0.77	37%	*	0.428	6%	*	0.098				
	<i>Micropterus dolomieu</i>	HUC-8	5%	*	0.12	0.07	0.12	-	-	-	-	-	-				
	<i>Micropterus salmoides</i>	HUC-4	3%	*	0.02	0.00	0.02	0%	ns	-	-	-	-				
Cottidae	<i>Cottus caroliniae</i>	HUC-8	66%	*	0.74	0.26	0.74	9%	*	0.772	15%	*	0.402				
	<i>Cottus hypselurus</i>	HUC-8	84%	ns	0.89	0.31	0.89	5%	ns	0.917	7%	*	0.442				
Fundulidae	<i>Fundulus catenatus</i>	HUC-8	36%	*	0.51	0.23	0.51	15%	*	0.516	16%	*	0.244				
	<i>Fundulus olivaceus</i>	HUC-8	18%	*	0.36	0.22	0.36	18%	*	0.372	21%	*	0.252				
Leuciscidae	<i>Campostoma anomalum</i>	HUC-8	53%	*	0.55	0.05	0.55	2%	*	0.680	7%	*	0.175				
	<i>Campostoma oligolepis</i>	HUC-8	6%	*	0.07	0.01	0.07	1%	ns	0.081	3%	*	0.036				
	<i>Chrosomus erythrogaster</i>	-	-	-	-	-	-	-	-	0.829	21%	*	0.548				
	<i>Cyprinella galactura</i>	HUC-8	7%	*	0.07	0.00	0.07	0%	ns	-	-	-	-				
	<i>Cyprinella whipplei</i>	HUC-8	14%	*	0.18	0.05	0.18	4%	*	0.202	7%	*	0.078				
	<i>Luxilus chrysocephalus</i>	HUC-8	14%	*	0.21	0.08	0.21	7%	*	0.266	10%	*	0.120				
	<i>Luxilus pilsbryi</i>	HUC-10	1%	ns	0.02	0.01	0.02	1%	-	-	-	-	-				
	<i>Luxilus zonatus</i>	HUC-10	15%	*	0.18	0.03	0.18	3%	*	0.199	10%	*	0.115				
	<i>Lythrurus umbratilis</i>	-	0%	ns	0.20	0.22	0.20	22%	*	-	-	-	-				
	<i>Notropis boops</i>	HUC-8	6%	*	0.09	0.03	0.09	3%	*	0.113	6%	*	0.059				
<i>Notropis nubilus</i>	HUC-4	10%	*	0.17	0.08	0.17	7%	*	0.172	1%	*	0.015					
<i>Notropis percobromus</i>	HUC-8	1%	*	0.01	0.01	0.01	1%	ns	-	-	-	-					
<i>Notropis telescopus</i>	HUC-8	33%	*	0.34	0.01	0.34	1%	*	0.436	3%	*	0.046					
<i>Pimephales notatus</i>	HUC-8	17%	*	0.44	0.32	0.44	26%	*	0.453	32%	*	0.372					
<i>Semotilus atromaculatus</i>	HUC-8	87%	ns	0.88	0.08	0.88	1%	*	0.934	2%	*	0.194					
Percidae	<i>Etheostoma blennioides</i>	HUC-8	61%	*	0.62	0.04	0.62	2%	*	0.686	2%	*	0.053				
	<i>Etheostoma caeruleum</i>	HUC-8	40%	*	0.44	0.06	0.44	3%	*	0.497	5%	*	0.093				
	<i>Etheostoma flabellare</i>	-	0%	ns	0.98	0.98	99%	*	0.977	3%	ns	0.580					
	<i>Etheostoma juliae</i>	HUC-8	34%	*	0.45	0.16	0.45	11%	*	0.478	12%	*	0.182				
<i>Etheostoma spectabile</i>	HUC-8	29%	*	0.38	0.14	0.38	10%	*	0.394	13%	*	0.181					
<i>Etheostoma zonale</i>	HUC-8	32%	*	0.34	0.02	0.34	2%	*	0.422	5%	*	0.074					
Poeciliidae	<i>Gambusia affinis</i>	HUC-4	7%	*	0.20	0.14	0.20	13%	ns	0.239	11%	*	0.123				

Abbreviations:  $\Phi_{ST}$ , genetic variation among sites across all groups;  $\Phi_{SC}$ , genetic variation among sites within a group; sig, the significance of the test (\* for <0.05 and ns for >0.05); Var, percent genetic variance explained.





**FIGURE 7** Neutral genetic variation was partitioned between four explanatory models for  $N = 31$  fish species sampled across the White River Basin (Ozark Mountains, USA). Partitioning was conducted separately for each species. The four models represent: (i) isolation by distance, the river network distance among individuals represented by spatial eigenvectors; (ii) isolation by barrier, represented by population structure coefficients among individuals; (iii) isolation by stream hierarchy, based on the hydrologic units (at four different hierarchical levels) in which an individual was collected; and (iv) isolation by environment, characterized by the environmental heterogeneity across sampling sites where individuals were collected. Total = the genetic variation explained by all four models combined. The "Pure" models represent the variation explained by each model after partialling out the variation explained by the other three models [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1111/mec.16806)]

species. Overall, stream fish genetic structure patterns indicated dispersal limited primarily among versus within river catchments.

## 4.1 | Drivers of isolation at the basin-wide scale

### 4.1.1 | Isolation by distance and river networks

IBD is expected when a genetic study's spatial extent is greater than individuals' average dispersal distance, that is, distance moved from natal habitat to breeding habitat. We examined species from seven families of fishes, most of which are small-bodied, highland species that presumably do not disperse great distances (Matthews & Robison, 1988). Indeed, significant IBD patterns were detected in 81% of the species in our study. However, the strength of the relationship was generally weak (Mantel  $r = 0.47$  and  $0.51$  for linearized  $F_{ST}$  and  $D$ , respectively).

While IBD may primarily explain genetic variation along a single stream or river, i.e., linear scale, it fails to incorporate the spatial

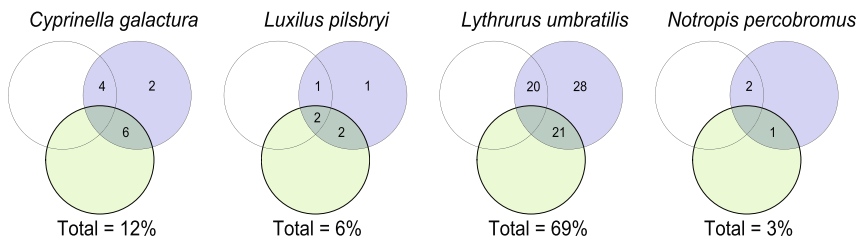
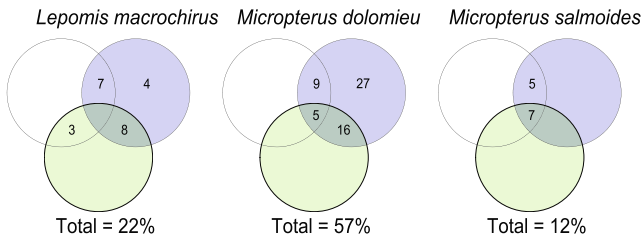
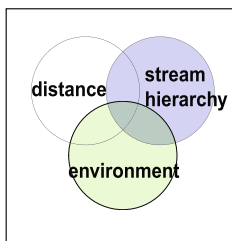
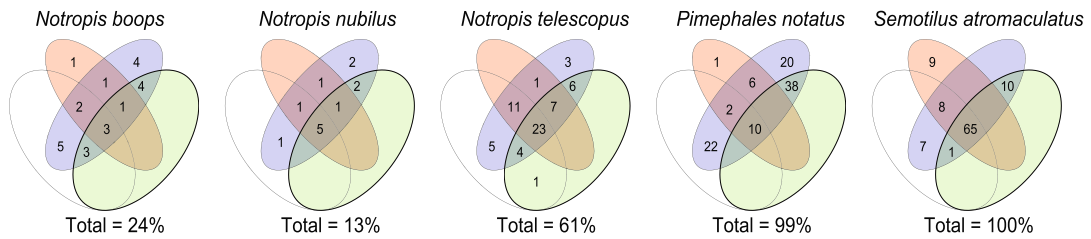
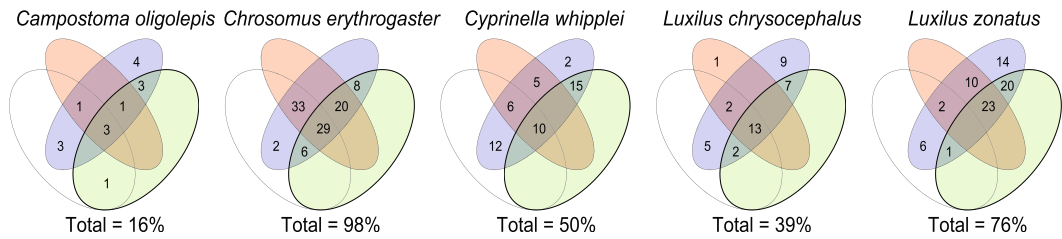
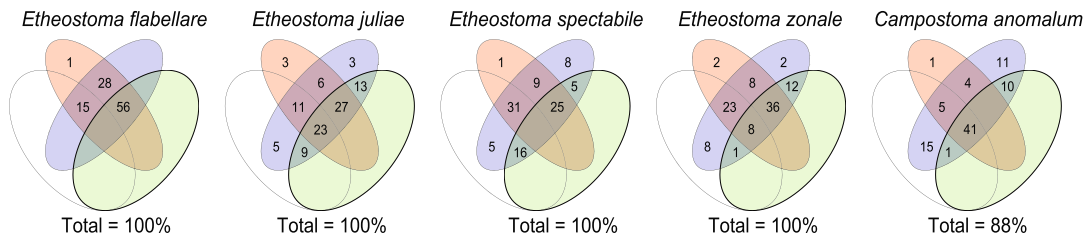
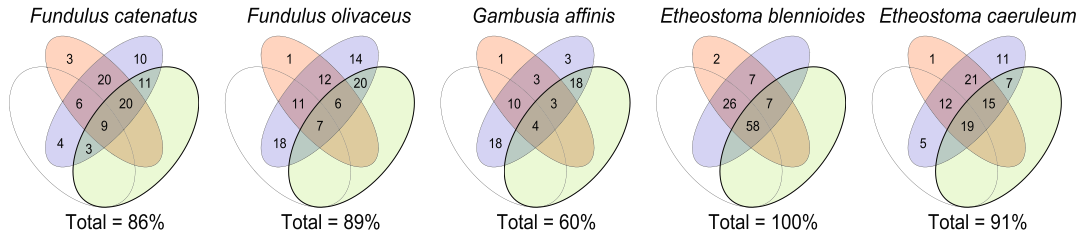
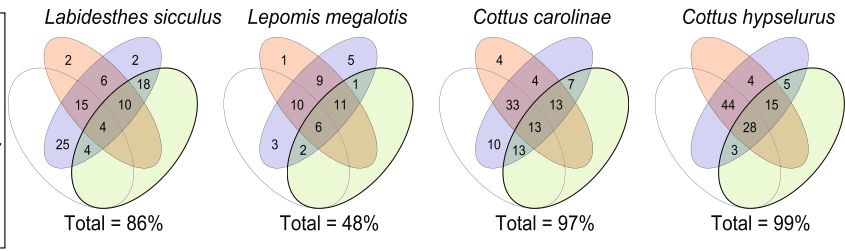
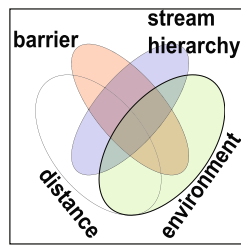
structure of riverine networks (Thomaz et al., 2016). Therefore, IBD may not be an appropriate general model for fish genetic structure at the network scale (Hopken et al., 2013). IBD plots for many species (Figure 4) showed high genetic divergence even among relatively proximate localities, with apparent clusters indicating discrete rather than continuous structure (Guillot et al., 2009). This evidence suggests that—at the network scale—a more nuanced pattern occurs, with high residual variation resulting. The failure of IBD to account for large amounts of variation in genetic divergence reflects additional resistance to dispersal, as caused by longitudinal changes in habitat characteristics such as slope, depth, volume, and predator composition. For example, two river reaches of equal length can have very different habitat matrices, and these can be more influential on gene flow than space alone (Guillot et al., 2009; Lowe et al., 2006; Ruiz-Gonzalez et al., 2015).

### 4.1.2 | Stream hierarchy model

Our results show that individual genetic variation is best explained by the stream hierarchy model (SHM) (Brauer et al., 2018; Hopken et al., 2013; Meffe & Vrijenhoek, 1988). In other words, the majority of variation explained by IBD, IBE, and IBB could be accounted for by IBH alone. This was corroborated via variation partitioning, in which IBD, IBE, and IBB models were redundant with IBH. A concordance of population structure with stream hierarchy yielded a similar percentage of among-site genetic variation, as explained by among-HUC and among-population groupings. In short, the variance explained by distance and environment was due to differences among HUC drainages. These results highlight the necessity of accounting for population structure prior to exploring the relationship between genotypes and environmental heterogeneity, for example, within genotype by environment frameworks (Lawson et al., 2020).

### 4.1.3 | Disentangling cumulative effects

Our analyses also revealed complex spatial patterns of genetic diversity. We evaluated competing isolation models using an approach that identified distance and barriers as putative drivers, with strong genetic divergence identified even across short geographical distances (Chan & Brown, 2020; Ruiz-Gonzalez et al., 2015). This interaction can confound analyses that incorporate either alone. For example, if sampling is clustered, discrete genetic groups can be spuriously inferred along an otherwise continuous gradient of genetic variation (Frantz et al., 2009). Furthermore, a continuous pattern can be erroneously extrapolated when the underlying reality is described by distinct clusters separated by geographic distance (Meirns, 2012). Here, we echo the importance of testing various hypotheses concerning genetic structure (Perez et al., 2018). Idiosyncrasies and complex interactions cannot be discerned by testing single models in isolation (e.g., discrete structure or IBD).



**FIGURE 8** Venn diagrams depict neutral genetic variation resulting from four models as applied to  $N = 31$  fish species sampled from the White River Basin (Ozark Mountains, USA). Models were based on: (i) isolation by distance, isolation by barrier, isolation by stream hierarchy, and isolation by environment. Values in the Venn diagrams are percent of genetic variance explained (i.e., rounded adjusted  $R^2$  values). Total variance explained is shown below each diagram. The bottom two rows show species that showed no discrete population structure (i.e., no isolation by barrier) and thus only three of the models were tested. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

## 4.2 | Drivers of variation within and among species

The species assayed herein display marked differences concerning dispersal capability (Shelley et al., 2021). Given this, we expected the degree of genetic structure to vary widely among species across our study region (Comte & Olden, 2018; Husemann et al., 2012; Pilger et al., 2017). Dispersal-related traits drive gene flow among localities and determine the spatial scale at which patterns of genetic structure emerge (Bohonak, 1999; Riginos et al., 2014). The physical structure of the river network then further modulates these patterns by dictating dispersal pathways of metapopulations and their colonization and extinction probabilities (Falke et al., 2012; Labonne et al., 2008; Fagan, 2002). These superimposed processes influence genetic divergence among distal populations (Thomaz et al., 2016; Chiu et al., 2020). Similar patterns emerge when analysing community diversity via species composition. Headwater streams tend to have very different communities due to dispersal limitations (Finn et al., 2011; Zbinden & Matthews, 2017; Zbinden, Geheber, Lehrter, & Matthews, 2022). The interaction between traits and the environment is an overarching influence that unites ecology and evolution.

Many species studied herein are small-bodied with aggregate distributions in upland and headwater streams (Robison & Buchanan, 2020). Thus, species-specific dispersal limitations, as imposed by unsuitable large riverine habitats (Radinger & Wolter, 2015; Schmidt & Schaefer, 2018), explain considerable variation in genetic structuring within the White River. Large rivers are hypothesized as inhospitable habitats to upland fishes (e.g., resources, depth, turbidity, substrates) and impose resistance to successful migration (e.g., higher discharge, greater density of large-bodied predators). These characteristics constrain migration and limit gene flow among basins that drain into large rivers (Fluker et al., 2014; Schmidt & Schaefer, 2018; Turner & Robison, 2006). The results are asymmetric gene flow and source-sink metapopulation dynamics, with susceptible species, those smaller and less tolerant, diverging most rapidly (Campbell Grant et al., 2007). Thus it is not surprising that the species for which we found no evidence of discrete genetic structure (i.e.,  $K = 1$ ) are also among the larger, more generalist species.

Life-history traits probably also play an important role. For example, those that directly influence effective population size (Nei & Tajima, 1981; Waples, 2022) may generate differences among species regarding the rate at which genetic differences arise due to genetic drift (Blanchet et al., 2020). Species with "slow" life histories, characterized by longer generations and delayed maturity, show an increased probability of local extirpation, inflating genetic drift concomitant with global extinction risk (Hutchings et al., 2012; Pearson et al., 2014; Chafin et al., 2019). Similar contingencies exist

for ecological traits, such as highly specialized trophic adaptations, narrow environmental tolerances, or any other following the same general mechanism of influence on dispersal or effective population sizes (Olden et al., 2008). Ecological traits are mirrored by morphology (Douglas & Matthews, 1992), underscoring an interaction of trait effects that are difficult to disentangle. Ultimately, intraspecific genetic divergence is driven by a combination of factors that influence population size, demographic history, and connectivity (Zbinden et al., 2022d). Clearly, these complex interactions among drivers require more comparative multispecies assessments as they shape genetic diversity and structure within and among species (microevolutionary scale) and, thus, ultimately lead to speciation and extinction (macroevolutionary scale).

## 4.3 | Disentangling historic and contemporary drivers

### 4.3.1 | Paleohydrology in the White River system

In this study, discrete population structure coincides with major topological divides within the White River stream network, such as a consistent east/west divide between Upper White and Black rivers, mirroring prior community composition studies (Matthews & Robison, 1988, 1998). Similar patterns were observed at smaller scales among drainages within the study region, as reported for White River crayfish (Fetzner Jr & DiStefano, 2008). While the Lower White and Black rivers are certainly contemporary large-river habitats, both would have been much larger pre-Pleistocene when together they represented the main channel of the Old Mississippi River (Mayden, 1988; Strange & Burr, 1997). This large-river habitat would have separated the eastern and western highland tributaries, with inhospitable habitat for upland species. Pronounced limitations regarding historic dispersal induced by the Old Mississippi could explain the greater isolation of the Little Red River and Black River tributary populations compared to those in the Upper White River. Here, additional work should incorporate coalescent perspectives (e.g., Oaks, 2019) that test the role of past geomorphic events in driving codivergence and codemographic patterns, such as the Pleistocene incursion by the Old Mississippi into the modern Black River channel.

### 4.3.2 | Contemporary drivers

Spatial discontinuities in genetic structure can also reveal contemporary barriers to migration/gene flow (Lee et al., 2018; Ruiz-Gonzalez

et al., 2015). The Upper White River dams (e.g., Norfolk, Bull Shoals, Table Rock, and Beaver dams) represent the most apparent anthropogenic barriers to gene flow. Dams elsewhere have demonstrated discrete populations above and below the structure (Roberts et al., 2013). However, observable genetic impacts can be limited due to the relatively short period these dams have been in place (Ruzich et al., 2019). Those on the White River were constructed between 1912 (Taneycomo Powersite Dam) and 1966 (Beaver Dam).

Our study was not explicitly designed to assess impoundment effects on diversity, nor were they directly tested. Nevertheless, evidence of discrete population structure has emerged, corresponding to the location of such dams. Four species showed discrete populations within the North Fork River above the Norfolk Dam: Southern Redbelly Dace (*Chrosomus erythrogaster*; Rafinesque, 1820); Yoke Darter (*Etheostoma juliae*; Meek, 1891); Northern Studfish (*Fundulus catenatus*; Storer, 1846); and Blackspotted Topminnow (*Fundulus olivaceus*; Storer, 1845; sites coloured magenta; Figure 6). One species, Orangethroat Darter (*Etheostoma spectabile*; Agassiz, 1854), showed a distinct population in the James River above Table Rock Dam (sites coloured gold; Figure 6). However, both North Fork and James rivers drain eight-digit HUC watersheds, which explains high amounts of genetic variation across the study region, regardless of dams. This highlights the importance of understanding “natural” network-wide patterns of genetic structure prior to deriving conclusions regarding anthropogenic barriers, particularly when they coincide with stream hierarchy. Differentiating dams as barriers versus stream hierarchy could be accomplished through divergence time estimates (Hansen et al., 2014) or by contrasting observed genetic differentiation with that expected in the face of a true obstacle for  $N$  generations (Prunier et al., 2020). That aspect, as it now stands, is beyond the scope of our current study.

## 5 | CONCLUSIONS

The multispecies comparative approach employed here revealed general patterns that could not have been discerned from a singular study of any one species. Additionally, the variability among species regarding intraspecific genetic structure provides specific information of interest that single-species studies cannot. While meta-analytic approaches have some potential, they are limited by confounding effects that stem from differences between studies, such as markers, sample sizes, environmental exigencies, and historical context. This necessitates a community-level approach within a study region. Further work aimed at modelling variables can lead to greater insight, ultimately improving our hypotheses regarding genetic diversity for which contemporary data are unavailable.

Importantly, our comparative approach supports the SHM as a general model for the genetic structure of lotic fish species and suggests that hydrologic units characterize regional genetic diversity quite well. Out of this result emerged the potential for HUC units to serve as a “rule of thumb” for riverine biodiversity conservation. None of the species evaluated herein were panmictic. Genetic

variation among HUCs was apparent despite limited evidence of discrete population or continuous structure for some species. Across a suite of commonly occurring fishes representing seven families, we identified greater intraspecific gene flow within than among basins/sub-basins. Therefore, fish populations within separate HUCs at the 8-digit+ level (e.g., HUC6, HUC4, HUC2) should be considered isolated until proven otherwise (Shelley et al., 2021).

As previously recognized, independent populations warrant independent management (Hopken et al., 2013). When migration is low or nonexistent, management of one population is unlikely to impact another. Genetic variation unique to hydrologic units could allow for adaptation to future environmental change, while on the other hand, isolation of populations could underscore elevated extirpation risks (Harrison et al., 2014). Furthermore, efforts to propagate populations via stocking or translocation should carefully assess the genetic landscape of the species in question, particularly before comingling diversity from different sub-basins (Meffe & Vrijenhoek, 1988). Such uninformed mixing of genetic stocks could promote outbreeding and the erosion of unique genetic diversity within river catchments. However, this must be weighed against the risks of local extirpation (Pavlova et al., 2017).

Given this study's general and comparative nature, we refrain from designating populations within species as potential management units (MUs). However, species showing high levels of genetic structure (Table 2) could be assessed individually for such designation, possibly requiring more fine-scaled, targeted sampling. Additional river/sub-basin-specific management efforts could also be justified, given the presence of unique populations across multiple species (Hopken et al., 2013). These consistent spatial modalities shared among multiple species point to areas of the river network that not only harbour unique genetic diversity but also play a key role in metapopulation viability and metacommunity dynamics (Fletcher et al., 2013). Here we specifically refer to: The Little Red, North Fork, Buffalo, Upper Black, Current, and Spring rivers. These may indeed represent evolutionarily significant catchments, and this insight underscores the potential for community-level genetic examination to elevate management to the ecosystem scale (Hanson et al., 2020; Paz-Vinas et al., 2018).

## AUTHOR CONTRIBUTIONS

Zachery D. Zbinden conceived the research with input from all authors. Specimen collection was done by Zachery D. Zbinden and Tyler K. Chafin. Zachery D. Zbinden did laboratory work, bioinformatics, data analysis, and manuscript drafting. All authors contributed to interpretation of results, formulating conclusions, and critically revising the manuscript. Marlis R. Douglas and Michael E. Douglas administered funding through their University of Arkansas endowments.

## ACKNOWLEDGEMENTS

We thank M. Flurry, M. George, T. Goodhart, K. Hollar, and M. Reed, who assisted with DNA extractions. The Arkansas High-Performance Computing Center provided analytical resources.

Funding was provided by the University of Arkansas Distinguished Doctoral Fellowship and Harry and Jo Leggett Chancellor's Fellowship (ZDZ), the Bruker Professorship in Life Sciences (MRD), the Twenty-First Century Chair in Global Change Biology (MED), and by an NSF Postdoctoral Research Fellowship in Biology (TKC) (DBI: 2010774). The findings, conclusions, and opinions expressed in this article represent those of the authors and do not necessarily represent the views of the NSF nor other affiliated or contributing organizations.

### CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

### DATA AVAILABILITY STATEMENT

Raw sequence reads have been deposited in the SRA (BioProject PRJNA809538). Species-level SNP panel alignments, metadata, and R code are available on Open Science Framework (White River Fish Genetic Structure; <http://osf.io/r8ytu>).

### BENEFIT-SHARING STATEMENT

No benefits to report.

### ORCID

Zachery D. Zbinden  <https://orcid.org/0000-0002-3709-224X>

Tyler K. Chafin  <https://orcid.org/0000-0001-8687-5905>

Michael E. Douglas  <https://orcid.org/0000-0001-9670-7825>

### REFERENCES

- Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12(1), 1–6. <https://doi.org/10.1186/1471-2105-12-246>
- Anderson, M. J., & Legendre, P. (1999). An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *Journal of Statistical Computation and Simulation*, 62(3), 271–303. <https://doi.org/10.1080/00949659908811936>
- Anthonyamy, W. J. B., Dreslik, M. J., Douglas, M. R., Thompson, D., Klut, G. M., Kuhns, A. R., Mauger, D., Kirk, D., Glowacki, G. A., Douglas, M. E., & Phillips, C. A. (2018). Population genetic evaluations within a co-distributed taxonomic group: A multispecies approach to conservation planning. *Animal Conservation*, 21(2), 137–147. <https://doi.org/10.1111/acv.12365>
- Avise, J. C. (1992). Molecular population structure and the biogeographic history of a regional fauna: A case history with lessons for conservation biology. *Oikos*, 63, 62–76. <https://doi.org/10.2307/3545516>
- Blanchet, F. G., Legendre, P., & Borcard, D. (2008). Forward selection of explanatory variables. *Ecology*, 89(9), 2623–2632. <https://doi.org/10.1890/07-0986.1>
- Blanchet, S., Prunier, J. G., & De Kort, H. (2017). Time to go bigger: Emerging patterns in macrogenetics. *Trends in Genetics*, 33(9), 579–580. <https://doi.org/10.1016/j.tig.2017.06.007>
- Blanchet, S., Prunier, J. G., Paz-Vinas, I., Saint-Pé, K., Rey, O., Raffard, A., Mathieu-Bégné, E., Loot, G., Fourtune, L., & Dubut, V. (2020). A river runs through it: The causes, consequences, and management of intraspecific diversity in river networks. *Evolutionary Applications*, 13(6), 1195–1213. <https://doi.org/10.1111/eva.12941>
- Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, 74(1), 21–45. <https://doi.org/10.1086/392950>
- Borcard, D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, 153(1–2), 51–68. [https://doi.org/10.1016/S0304-3800\(01\)00501-4](https://doi.org/10.1016/S0304-3800(01)00501-4)
- Borcard, D., Legendre, P., & Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology*, 73(3), 1045–1055. <https://doi.org/10.2307/1940179>
- Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution*, 67(11), 3258–3273. <https://doi.org/10.1111/evo.12193>
- Brauer, C. J., Unmack, P. J., Smith, S., Bernatchez, L., & Beheregaray, L. B. (2018). On the roles of landscape heterogeneity and environmental variation in determining population genomic structure in a dendritic system. *Molecular Ecology*, 27(17), 3484–3497. <https://doi.org/10.1111/mec.14808>
- Campbell Grant, E. H., Lowe, W. H., & Fagan, W. F. (2007). Living in the branches: Population dynamics and ecological processes in dendritic networks. *Ecology Letters*, 10(2), 165–175. <https://doi.org/10.1111/j.1461-0248.2006.01007.x>
- Capblancq, T., & Forester, B. R. (2021). Redundancy analysis: A Swiss Army knife for landscape genomics. *Methods in Ecology and Evolution*, 12(12), 2298–2309. <https://doi.org/10.1111/2041-210X.13722>
- Chafin, T. K., Douglas, M. R., Martin, B. T., & Douglas, M. E. (2019). Hybridization drives genetic erosion in sympatric desert fishes of western North America. *Heredity*, 123(6), 759–773. <https://doi.org/10.1038/s41437-019-0259-2>
- Chan, K. O., & Brown, R. M. (2020). Elucidating the drivers of genetic differentiation in Malaysian torrent frogs (Anura: Ranidae: Amolops): A landscape genomics approach. *Zoological Journal of the Linnean Society*, 190(1), 65–78. <https://doi.org/10.1093/zoolinnean/zlz151>
- Chiu, M. C., Li, B., Nukazawa, K., Resh, V. H., Carvajal, T., & Watanabe, K. (2020). Branching networks can have opposing influences on genetic variation in riverine metapopulations. *Diversity and Distributions*, 26(12), 1813–1824. <https://doi.org/10.1111/ddi.13160>
- Comte, L., & Olden, J. D. (2018). Fish dispersal in flowing waters: A synthesis of movement- and genetic-based studies. *Fish and Fisheries*, 19, 1063–1077. <https://doi.org/10.1111/faf.12312>
- Coomes, K. R., & Wang, M. (2019). *Package 'PCDimension'*. Retrieved from: <https://cran.r-project.org/package=PCDimension>
- Cushman, S. A., McKelvey, K. S., Hayden, J., & Schwartz, M. K. (2006). Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *American Naturalist*, 168(4), 486–499. <https://doi.org/10.1086/506976>
- da Fonseca, R. R., Albrechtsen, A., Themudo, G. E., Ramos-Madrugal, J., Sibbesen, J. A., Maretty, L., Zepeda-Mendoza, M. L., Campos, P. F., Heller, R., & Pereira, R. J. (2016). Next-generation biology: Sequencing and data analysis approaches for non-model organisms. *Marine Genomics*, 30, 3–13. <https://doi.org/10.1016/j.margen.2016.04.012>
- Davis, C. D., Epps, C. W., Flitcroft, R. L., & Banks, M. A. (2018). Refining and defining riverscape genetics: How rivers influence population genetic structure. *Wiley Interdisciplinary Reviews: Water*, 5(2), e1269. <https://doi.org/10.1002/wat2.1269>
- De Kort, H., Prunier, J. G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V. M., & Blanchet, S. (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. *Nature Communications*, 12(1), 1–11. <https://doi.org/10.1038/s41467-021-20958-2>
- DeWoody, J. A., Harder, A. M., Mathur, S., & Willoughby, J. R. (2021). The long-standing significance of genetic diversity in conservation. *Molecular Ecology*, 30(17), 4147–4154. <https://doi.org/10.1111/mec.16051>
- Douglas, M. E., & Matthews, W. J. (1992). Does morphology predict ecology? Hypothesis testing within a freshwater stream fish assemblage. *Oikos*, 65(2), 213–224. <https://doi.org/10.2307/3545012>



- Douglas, M. R., Anthonyamy, W. J. B., Mussmann, S. M., Davis, M. A., Louis, W., & Douglas, M. E. (2020). Multi-targeted management of upland game birds at the agroecosystem interface in mid-western North America. *PLoS One*, 15(4), e0230735. <https://doi.org/10.1371/journal.pone.0230735>
- Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque, G., Legendre, P., & Wagner, H. (2020). *R package "adespatial": Multivariate multiscale spatial analysis*. Retrieved from: <https://cran.r-project.org/package=adespatial>
- Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, 196(3–4), 483–493. <https://doi.org/10.1016/j.ecolmodel.2006.02.015>
- Eaton, D. A. R., & Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, 36(8), 2592–2594. <https://doi.org/10.1093/bioinformatics/btz966>
- Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131(2), 479–491. <https://doi.org/10.1093/genetics/131.2.479>
- Fagan, W. F. (2002). Connectivity, fragmentation, and extinction risk in dendritic metapopulations. *Ecology*, 83(12), 3243–3249. [https://doi.org/10.1890/0012-9658\(2002\)083\[3243:CFAERI\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[3243:CFAERI]2.0.CO;2)
- Falke, J. A., Bailey, L. L., Fausch, K. D., & Bestgen, K. R. (2012). Colonization and extinction in dynamic habitats: An occupancy approach for a Great Plains stream fish assemblage. *Ecology*, 93(4), 858–867. <https://doi.org/10.1890/11-1515.1>
- Fetzner, J. W., Jr., & DiStefano, R. J. (2008). Population genetics of an imperiled crayfish from the White River drainage of Missouri, USA. *Freshwater Crayfish*, 16, 131–146.
- Finn, D. S., Bonada, N., Múrria, C., & Hughes, J. M. (2011). Small but mighty: Headwaters are vital to stream network biodiversity at two levels of organization. *Journal of the North American Benthological Society*, 30(4), 963–980. <https://doi.org/10.1899/11-012.1>
- Fletcher, R. J., Revell, A., Reichert, B. E., Kitchens, W. M., Dixon, J. D., & Austin, J. D. (2013). Network modularity reveals critical scales for connectivity in ecology and evolution. *Nature Communications*, 4(1), 1–7. <https://doi.org/10.1038/ncomms3572>
- Fluker, B. L., Kuhajda, B. R., & Harris, P. M. (2014). The influence of life-history strategy on genetic differentiation and lineage divergence in darters (Percidae: Etheostominae). *Evolution*, 68(11), 3199–3216. <https://doi.org/10.1111/evo.12505>
- Fortuna, M. A., Albaladejo, R. G., Fernández, L., Aparicio, A., & Bascompte, J. (2009). Networks of spatial genetic variation across species. *Proceedings of the National Academy of Sciences*, 106(45), 19044–19049. <https://doi.org/10.1073/pnas.0907704106>
- Frantz, A. C., Cellina, S., Krier, A., Schley, L., & Burke, T. (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: Clusters or isolation by distance? *Journal of Applied Ecology*, 46(2), 493–505. <https://doi.org/10.1111/j.1365-2664.2008.01606.x>
- Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925–929. <https://doi.org/10.1111/2041-210X.12382>
- Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, 196(4), 973–983. <https://doi.org/10.1534/genetics.113.160572>
- Goslee, S., & Urban, D. (2020). *Ecodist: Dissimilarity-based functions for ecological analysis*. Retrieved from: <https://cran.r-project.org/package=ecodist>
- Gosselin, T. (2020). *R package radiator: RADseq data exploration, manipulation and visualization using R*. Retrieved from: <https://thierrygoselin.github.io/radiator/>
- Goudet, J., Jombart, T., Kamvar, Z. N., Archer, E., & Hardy, O. (2017). Hierfstat: Estimation and tests of hierarchical F-statistics. Retrieved from: <https://cran.r-project.org/package=hierfstat>
- Guillot, G., Leblois, R., Coulon, A., & Frantz, A. C. (2009). Statistical methods in spatial genetics. *Molecular Ecology*, 18(23), 4734–4756. <https://doi.org/10.1111/j.1365-294X.2009.04410.x>
- Hansen, M. M., Limborg, M. T., Ferchaud, A. L., & Pujolar, J. M. (2014). The effects of medieval dams on genetic divergence and demographic history in brown trout populations. *BMC Evolutionary Biology*, 14(1), 1–14. <https://doi.org/10.1186/1471-2148-14-122>
- Hanson, J. O., Marques, A., Verissimo, A., Camacho-Sanchez, M., Velo-Antón, G., Martínez-Solano, Í., & Carvalho, S. B. (2020). Conservation planning for adaptive and neutral evolutionary processes. *Journal of Applied Ecology*, 57(11), 2159–2169. <https://doi.org/10.1111/1365-2664.13718>
- Harrison, K. A., Pavlova, A., Telonis-Scott, M., & Sunnucks, P. (2014). Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, 7(9), 1008–1025. <https://doi.org/10.1111/eva.12149>
- Holderregger, R., Kamm, U., & Gugerli, F. (2006). Adaptive vs. neutral genetic diversity: Implications for landscape genetics. *Landscape Ecology*, 21(6), 797–807. <https://doi.org/10.1007/s10980-005-5245-9>
- Hopken, M. W., Douglas, M. R., & Douglas, M. E. (2013). Stream hierarchy defines riverscape genetics of a north American desert fish. *Molecular Ecology*, 22(4), 956–971. <https://doi.org/10.1111/mec.12156>
- Hotelling, S., Giersch, J. J., Finn, D. S., Tronstad, L. M., Jordan, S., Serpa, L. E., Call, R. G., Muhlfeld, C. C., & Weisrock, D. W. (2019). Congruent population genetic structure but differing depths of divergence for three alpine stoneflies with similar ecology and geographic distributions. *Freshwater Biology*, 64(2), 335–347. <https://doi.org/10.1111/fwb.13223>
- Huber, P. R., Greco, S. E., & Thorne, J. H. (2010). Spatial scale effects on conservation network design: Trade-offs and omissions in regional versus local scale planning. *Landscape Ecology*, 25(5), 683–695. <https://doi.org/10.1007/s10980-010-9447-4>
- Hubert, M., Reynkens, T., Schmitt, E., & Verdonck, T. (2016). Sparse PCA for high-dimensional data with outliers. *Technometrics*, 58(4), 424–434. <https://doi.org/10.1080/00401706.2015.1093962>
- Hubert, M., Rousseeuw, P. J., & Vanden Branden, K. (2005). ROBPCA: A new approach to robust principal component analysis. *Technometrics*, 47(1), 64–79. <https://doi.org/10.1198/00401700400000563>
- 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>
- Husemann, M., Ray, J. W., King, R. S., Hooser, E. A., & Danley, P. D. (2012). Comparative biogeography reveals differences in population genetic structure of five species of stream fishes. *Biological Journal of the Linnean Society*, 107(4), 867–885. <https://doi.org/10.1111/j.1095-8312.2012.01973.x>
- Hutchings, J. A., Myers, R. A., García, V. B., Lucifora, L. O., & Kuparinen, A. (2012). Life-history correlates of extinction risk and recovery potential. *Ecological Applications*, 22(4), 1061–1067. <https://doi.org/10.1890/11-1313.1>
- Hutchison, D. W., & Templeton, A. R. (1999). Correlation of pairwise genetic and geographic distance measures: Inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution*, 53(6), 1898–1914. <https://doi.org/10.1111/j.1558-5646.1999.tb04571.x>
- Jones, O. R., & Wang, J. (2012). A comparison of four methods for detecting weak genetic structure from marker data. *Ecology and Evolution*, 2(5), 1048–1055. <https://doi.org/10.1002/ece3.237>

- Jorgensen, D. J. (1993). *Paleohydrology of the Central United States*. U.S. Geological Survey Bulletin 1989-D, 1–32. <https://doi.org/10.3133/b1989D>
- Jost, L. (2008). GST and its relatives do not measure differentiation. *Molecular Ecology*, 17(18), 4015–4026. <https://doi.org/10.1111/j.1365-294X.2008.03887.x>
- Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, 6, 208. <https://doi.org/10.3389/fgene.2015.00208>
- Labonne, J., Ravigné, V., Parisi, B., & Gaucherel, C. (2008). Linking dendritic network structures to population demogenetics: The downside of connectivity. *Oikos*, 117(10), 1479–1490. <https://doi.org/10.1111/j.0030-1299.2008.16976.x>
- Laikre, L. (2010). Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics*, 11(2), 349–354. <https://doi.org/10.1007/s10592-009-0037-4>
- Lawson, D. J., Davies, N. M., Haworth, S., Ashraf, B., Howe, L., Crawford, A., Hemani, G., Smith, G. D., & Timpson, N. J. (2020). Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity? *Human Genetics*, 139, 23–41. <https://doi.org/10.1007/s00439-019-02014-8>
- Lee, S. R., Jo, Y. S., Park, C. H., Friedman, J. M., & Olson, M. S. (2018). Population genomic analysis suggests strong influence of river network on spatial distribution of genetic variation in invasive saltcedar across the southwestern United States. *Molecular Ecology*, 27(3), 636–646. <https://doi.org/10.1111/mec.14468>
- Legendre, P., & Fortin, M. J. (2010). Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*, 10(5), 831–844. <https://doi.org/10.1111/j.1755-0998.2010.02866.x>
- Leonard, P. B., Baldwin, R. F., & Hanks, R. D. (2017). Landscape-scale conservation design across biotic realms: Sequential integration of aquatic and terrestrial landscapes. *Scientific Reports*, 7(1), 1–12. <https://doi.org/10.1038/s41598-017-15304-w>
- Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., & Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Scientific Data*, 6(1), 1–15. <https://doi.org/10.1038/s41597-019-0300-6>
- Lowe, W. H., Likens, G. E., McPeck, M. A., & Buso, D. C. (2006). Linking direct and indirect data on dispersal: Isolation by slope in a headwater stream salamander. *Ecology*, 87(2), 334–339. <https://doi.org/10.1890/05-0232>
- Mantel, N., & Valand, R. S. (1970). A technique of nonparametric multivariate analysis. *Biometrics*, 26(3), 547. <https://doi.org/10.2307/2529108>
- Margules, C. R., & Pressey, R. L. (2000). Systematic conservation planning. *Nature*, 405, 243–253. <https://doi.org/10.1038/35012251>
- Matthews, W. J. (1998). *Patterns in freshwater fish ecology*. Chapman and Hall.
- Matthews, W. J., & Robison, H. W. (1988). The distribution of the fishes of Arkansas: A multivariate analysis. *Copeia*, 1988(2), 358–374. <https://doi.org/10.2307/1445876>
- Matthews, W. J., & Robison, H. W. (1998). Influence of drainage connectivity, drainage area and regional species richness on fishes of the interior highlands in Arkansas. *American Midland Naturalist*, 139(1), 1–19. [https://doi.org/10.1674/0003-0031\(1998\)139\[0001:IODCD\]2.0.CO;2](https://doi.org/10.1674/0003-0031(1998)139[0001:IODCD]2.0.CO;2)
- 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>
- McRae, B. H. (2006). Isolation by resistance. *Evolution*, 60(8), 1551–1561. <https://doi.org/10.1111/j.0014-3820.2006.tb00500.x>
- Meffe, G. K., & Vrijenhoek, R. C. (1988). Conservation genetics in the management of desert fishes. *Conservation Biology*, 2(2), 157–169. <https://doi.org/10.1111/j.1523-1739.1988.tb00167.x>
- Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, 21(12), 2839–2846. <https://doi.org/10.1111/j.1365-294X.2012.05578.x>
- Meirmans, P. G. (2015). Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, 24(13), 3223–3231. <https://doi.org/10.1111/mec.13243>
- Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, 11(1), 5–18. <https://doi.org/10.1111/j.1755-0998.2010.02927.x>
- Naimi, B. (2013). *R package usdm: Uncertainty analysis for species distribution models*. Retrieved from: <https://cran.r-project.org/package=usdm>
- Nazareno, A. G., Bemmels, J. B., Dick, C. W., & Lohmann, L. G. (2017). Minimum sample sizes for population genomics: An empirical study from an Amazonian plant species. *Molecular Ecology Resources*, 17(6), 1136–1147. <https://doi.org/10.1111/1755-0998.12654>
- Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences*, 70, 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321>
- Nei, M., & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics*, 98(3), 625–640. <https://doi.org/10.1093/genetics/98.3.625>
- Oaks, J. R. (2019). Full Bayesian comparative phylogeography from genomic data. *Systematic Biology*, 68(3), 371–395. <https://doi.org/10.1093/sysbio/syy063>
- Oksanen, A. J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). R package vegan: Community ecology package. Retrieved from: <https://cran.r-project.org/package=vegan>
- Olden, J. D., Poff, N. L., & Bestgen, K. R. (2008). Trait synergisms and the rarity, extirpation, and extinction risk of desert fishes. *Ecology*, 89(3), 847–856. <https://doi.org/10.1890/06-1864.1>
- Orsini, L., Vanoverbeke, J., Swillen, I., Mergeay, J., & De Meester, L. (2013). Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, 22(24), 5983–5999. <https://doi.org/10.1111/mec.12561>
- Paradis, E. (2010). Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, 26(3), 419–420. <https://doi.org/10.1093/bioinformatics/btp696>
- Pavlova, A., Beheregaray, L. B., Coleman, R., Gilligan, D., Harrison, K. A., Ingram, B. A., Kearns, J., Lamb, A. M., Lintermans, M., Lyon, J., Nguyen, T. T. T., Sasaki, M., Tonkin, Z., Yen, J. D. L., & Sunnucks, P. (2017). Severe consequences of habitat fragmentation on genetic diversity of an endangered Australian freshwater fish: A call for assisted gene flow. *Evolutionary Applications*, 10(6), 531–550. <https://doi.org/10.1111/eva.12484>
- Paz-Vinas, I., & Blanchet, S. (2015). Dendritic connectivity shapes spatial patterns of genetic diversity: A simulation-based study. *Journal of Evolutionary Biology*, 28(4), 986–994. <https://doi.org/10.1111/jeb.12626>
- Paz-Vinas, I., Loot, G., Hermoso, V., Veyssi ere, C., Poulet, N., Grenouillet, G., & Blanchet, S. (2018). Systematic conservation planning for intraspecific genetic diversity. *Proceedings of the Royal Society B: Biological Sciences*, 285(1877), 20172746. <https://doi.org/10.1098/rspb.2017.2746>
- Paz-Vinas, I., Loot, G., Stevens, V. M., & Blanchet, S. (2015). Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Molecular Ecology*, 24(18), 4586–4604. <https://doi.org/10.1111/mec.13345>

- Pearson, R. G., Stanton, J. C., Shoemaker, K. T., Aiello-Lammens, M. E., Ersts, P. J., Horning, N., Fordham, D. A., Raxworthy, C. J., Ryu, H. Y., McNeese, J., & Akçakaya, H. R. (2014). Life history and spatial traits predict extinction risk due to climate change. *Nature Climate Change*, 4(3), 217–221. <https://doi.org/10.1038/nclimate2113>
- Peres-Neto, P. R., Jackson, D. A., & Somers, K. M. (2005). How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis*, 49(4), 974–997. <https://doi.org/10.1016/j.csda.2004.06.015>
- Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: Estimation and comparison of fractions. *Ecology*, 87(10), 2614–2625. [https://doi.org/10.1890/0012-9658\(2006\)87\[2614:VPOSDM\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2)
- Perez, M. F., Franco, F. F., Bombonato, J. R., Bonatelli, I. A. S., Khan, G., Romeiro-Brito, M., Fegies, A. C., Ribeiro, P. M., Silva, G. A. R., & Moraes, E. M. (2018). Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Diversity and Distributions*, 24(12), 1883–1889. <https://doi.org/10.1111/ddi.12816>
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Peterson, E. E., Ver Hoef, J. M., Isaak, D. J., Falke, J. A., Fortin, M. J., Jordan, C. E., McNyset, K., Monestiez, P., Ruesch, A. S., Sengupta, A., Som, N., Steel, E. A., Theobald, D. M., Torgersen, C. E., & Wenger, S. J. (2013). Modelling dendritic ecological networks in space: An integrated network perspective. *Ecology Letters*, 16(5), 707–719. <https://doi.org/10.1111/ele.12084>
- Pilger, T. J., Gido, K. B., Propst, D. L., Whitney, J. E., & Turner, T. F. (2017). River network architecture, genetic effective size and distributional patterns predict differences in genetic structure across species in a dryland stream fish community. *Molecular Ecology*, 26(10), 2687–2697. <https://doi.org/10.1111/mec.14079>
- Prunier, J. G., Poesy, C., Dubut, V., Veyssi re, C., Loot, G., Poulet, N., & Blanchet, S. (2020). Quantifying the individual impact of artificial barriers in freshwaters: A standardized and absolute genetic index of fragmentation. *Evolutionary Applications*, 13(10), 2566–2581. <https://doi.org/10.1111/eva.13044>
- Radinger, J., & Wolter, C. (2015). Disentangling the effects of habitat suitability, dispersal, and fragmentation on the distribution of river fishes. *Ecological Applications*, 25(4), 914–927. <https://doi.org/10.1890/14-0422.1>
- Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *Sankhy : The Indian Journal of Statistics Series A*, 26, 329–358. <https://www.jstor.org/stable/25049339>
- Riginos, C., Buckley, Y. M., Blomberg, S. P., & Tremblay, E. A. (2014). Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *The American Naturalist*, 184(1), 52–64. <https://doi.org/10.1086/676505>
- Roberts, J. H., Angermeier, P. L., & Hallerman, E. M. (2013). Distance, dams and drift: What structures populations of an endangered, benthic stream fish? *Freshwater Biology*, 58(10), 2050–2064. <https://doi.org/10.1111/fwb.12190>
- Robins, C. R., & Robison, H. W. (1985). *Cottus hypselurus*, a new cottid fish from the ozark uplands, Arkansas and Missouri. *American Midland Naturalist*, 114(2), 360. <https://doi.org/10.2307/2425610>
- Robison, H. W., & Buchanan, T. M. (2020). *Fishes of Arkansas* (2nd ed.). University of Arkansas Press.
- Rosenberg, N. A., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J. K., & Feldman, M. W. (2005). Clines, clusters, and the effect of study design on the inference of human population structure. *PLoS Genetics*, 1(6), 660–671. <https://doi.org/10.1371/journal.pgen.0010070>
- Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145(4), 1219–1228. <https://doi.org/10.1006/bjil.1996.0127>
- Ruiz-Gonzalez, A., Cushman, S. A., Madeira, M. J., Randi, E., & G mez-Moliner, B. J. (2015). Isolation by distance, resistance and/or clusters? Lessons learned from a forest-dwelling carnivore inhabiting a heterogeneous landscape. *Molecular Ecology*, 24(20), 5110–5129. <https://doi.org/10.1111/mec.13392>
- Ruzich, J., Turnquist, K., Nye, N., Rowe, D., & Larson, W. A. (2019). Isolation by a hydroelectric dam induces minimal impacts on genetic diversity and population structure in six fish species. *Conservation Genetics*, 20(6), 1421–1436. <https://doi.org/10.1007/s10592-019-01220-1>
- Schmidt, B. V., & Schaefer, J. (2018). Ecological and landscape effects on genetic distance in an assemblage of headwater fishes. *Ecology of Freshwater Fish*, 27(2), 617–631. <https://doi.org/10.1111/eff.12375>
- Shelley, J. J., Holland, O. J., Swearer, S. E., Dempster, T., Le, M. C., Craig, F., & Miller, A. D. (2021). Landscape context and dispersal ability as determinants of population genetic structure in freshwater fishes. *Freshwater Biology*, 67(2), 338–352. <https://doi.org/10.1111/fwb.13844>
- Singhal, S., Huang, H., Grundler, M. R., March n-Rivadeneira, M. R., Holmes, I., Title, P. O., Donnellan, S. C., & Rabosky, D. L. (2018). Does population structure predict the rate of speciation? A comparative test across Australia's most diverse vertebrate radiation. *The American Naturalist*, 192(4), 432–447. <https://doi.org/10.1086/699515>
- 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>
- Thomaz, A. T., Christie, M. R., & Knowles, L. L. (2016). The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution*, 70(3), 731–739. <https://doi.org/10.1111/evo.12883>
- Turner, T. F., & Robison, H. W. (2006). Genetic diversity of the Caddo Madtom, *Noturus taylori*, with comments on factors that promote genetic divergence in fishes endemic to the Ouachita highlands. *Southwestern Naturalist*, 51(3), 338–345. [https://doi.org/10.1894/0038-4909\(2006\)51\[338:GDOTCM\]2.0.CO;2](https://doi.org/10.1894/0038-4909(2006)51[338:GDOTCM]2.0.CO;2)
- Tyers, M. (2017). *Riverdist: River network distance computation and applications*. Retrieved from: <https://cran.r-project.org/package=riverdist>
- USGS. (2021). *Watershed Boundary Dataset*. Retrieved from: <https://www.usgs.gov/national-hydrography/watershed-boundary-dataset>. Accessed: 1 December 2021.
- USGS, & USDA-NRCS. (2013). Federal standards and procedures for the National Watershed Boundary Dataset (WBD). (Techniques and Methods 11-A3, 4th ed.). U.S. Geological Survey and U.S. Department of Agriculture, Natural Resources Conservation Service. <https://pubs.usgs.gov/tm/11/a3/>
- Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, 23(23), 5649–5662. <https://doi.org/10.1111/mec.12938>
- Waples, R. S. (2022). What is  $N_e$ , anyway? *Journal of Heredity*, 113(4), 371–379. <https://doi.org/10.1093/jhered/esac023>
- Warren, M. L., Burr, B. M., Walsh, S. J., Bart, H. L., Cashner, R. C., Etnier, D. A., Freeman, B. J., Kuhajda, B. R., Mayden, R. L., Robison, H. W., Ross, S. T., & Starnes, W. C. (2000). Diversity, distribution, and conservation status of the native freshwater fishes of the southern United States. *Fisheries*, 25(10), 7–31. [https://doi.org/10.1577/1548-8446\(2000\)025<0007:ddacso>2.0.co;2](https://doi.org/10.1577/1548-8446(2000)025<0007:ddacso>2.0.co;2)
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>

- White, S. L., Hanks, E. M., & Wagner, T. (2020). A novel quantitative framework for riverscape genetics. *Ecological Applications*, 30(7), 1–17. <https://doi.org/10.1002/eap.2147>
- Willing, E. M., Dreyer, C., & van Oosterhout, C. (2012). Estimates of genetic differentiation measured by Fst do not necessarily require large sample sizes when using many snp markers. *PLoS One*, 7(8), e42649. <https://doi.org/10.1371/journal.pone.0042649>
- Winter, D. J. (2012). MMOD: An R library for the calculation of population differentiation statistics. *Molecular Ecology Resources*, 12(6), 1158–1160. <https://doi.org/10.1111/j.1755-0998.2012.03174.x>
- Wright, S. (1943). Isolation by distance. *Genetics*, 28(2), 114–138.
- Xuereb, A., Benestan, L., Normandeau, É., Daigle, R. M., Curtis, J. M., Bernatchez, L., & Fortin, M. J. (2018). Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RAD seq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Molecular Ecology*, 27(10), 2347–2364. <https://doi.org/10.1111/mec.14589>
- Xuereb, A., D'Aloia, C. C., Andrello, M., Bernatchez, L., & Fortin, M.-J. (2021). Incorporating putatively neutral and adaptive genomic data into marine conservation planning. *Conservation Biology*, 35(3), 909–920. <https://doi.org/10.1111/cobi.13609>
- Zbinden, Z. D., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2022a). Ozark fish community genomics: ddRAD seq. *NCBI sequence read archive*. BioProject: PRJNA809538. [dataset].
- Zbinden, Z. D., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2022b). A community genomics approach to natural hybridization. *bioRxiv*. <https://doi.org/10.1101/2022.10.26.513841>
- Zbinden, Z. D., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2022c). White River fish genetic structure. *Open Science Framework*. <https://doi.org/10.17605/OSF.IO/R8YTU> [dataset].
- Zbinden, Z. D., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2022d). Trait-based predictors of genomic diversity within freshwater fish communities of the Ozarks. *bioRxiv*. <https://doi.org/10.1101/2022.10.26.513843>
- Zbinden, Z. D., Geheber, A. D., Lehrter, R. J., & Matthews, W. J. (2022). Multifaceted assessment of stream fish alpha and beta diversity using spatial models. *Hydrobiologia*, 849, 1795–1820. <https://doi.org/10.1007/s10750-022-04824-w>
- Zbinden, Z. D., Geheber, A. D., Matthews, W. J., & Marsh-Matthews, E. (2022). Fish communities, species of greatest conservation need, and protected areas in southeastern Oklahoma, 2014–2016. *Proceedings of the Oklahoma Academy of Science*, 101, 14–32.
- Zbinden, Z. D., & Matthews, W. J. (2017). Beta diversity of stream fish assemblages: Partitioning variation between spatial and environmental factors. *Freshwater Biology*, 62(8), 1460–1471. <https://doi.org/10.1111/fwb.12960>

## SUPPORTING INFORMATION

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

**How to cite this article:** Zbinden, Z. D., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2023). Riverscape community genomics: A comparative analytical approach to identify common drivers of spatial structure. *Molecular Ecology*, 32, 6743–6765. <https://doi.org/10.1111/mec.16806>