ORIGINAL ARTICLE

WILEY Freshwater Biology

A genomic assessment of population structure and gene flow in an aquatic salamander identifies the roles of spatial scale, barriers, and river architecture

Mason O. Murphy[1] [iD] | Kara S. Jones[1] | Steven J. Price[2] | David W. Weisrock[1]

[1]Department of Biology, University of Kentucky, Lexington, KY, USA

[2]Department of Forestry and Natural Resources, University of Kentucky, Lexington, KY, USA

Correspondence
David W. Weisrock, Department of Biology, University of Kentucky, Lexington, KY, USA.
Email: david.weisrock@uky.edu

Present address
Mason O. Murphy, Department of Biology, Miami University, Oxford, OH, 45056, USA.

Abstract

1. Population structure and gene flow of species in lotic environments can be constrained by river network architecture, species life history and heterogeneous local barriers. Identifying the factors that influence population structure and gene flow, especially in species limited to movement within a river network, is vital for understanding the evolutionary and demographic history of a species.

2. We explored population structure and gene flow for a fully aquatic salamander, the common mudpuppy (*Necturus maculosus*), in Kentucky (USA) using genomic data. We examined population structure using both parametric and nonparametric methods, and we tested for a history of lineage divergence among identified genetic clusters. We quantified the partitioning of genetic variation at different hierarchical levels, and we tested for signatures of isolation by distance. Additionally, we used coalescent-based model selection to identify a best-fit model of gene flow between our three sampled basins.

3. We found the greatest support for population structure between the Kentucky River basin and the combined Licking and Kinniconick basins, with further subdivision within both the Kentucky and Licking River basins. However, we found no evidence for a history of lineage divergence among these structured units. The movement of *N. maculosus* is constrained by the lotic network architecture, which likely drives the evolution of this hierarchical population structure, with increasing differentiation between sites nested in river basins, and even greater differentiation between basins. However, we also found evidence for population structure not explained by river architecture, with an isolated population embedded within the Kentucky River basin.

4. This study demonstrates the heterogeneity in population structure that can evolve in aquatic species occupying lotic systems and illustrates the potential for genomic data to disentangle these complex patterns.

KEYWORDS
Necturus maculosus, Kentucky, lotic network, model testing, population genetics

1 | INTRODUCTION

Population structure varies across a spectrum of divergence, ranging from panmixia to complete isolation of populations (Hutchison & Templeton, 1999; Wright, 1949). Patterns of population structure found across the range of a species are influenced by the interaction between intrinsic factors, such as life history, and extrinsic factors, such as landscape architecture and dispersal barriers (Coulon et al., 2006; Finn, Blouin, & Lytle, 2007). While intrinsic factors are expected to affect a species similarly throughout its range, the effect of extrinsic factors on population structure can vary because local landscapes offer different levels of resistance to gene flow (Zeller, McGarigal, & Whiteley, 2012).

Extrinsic factors are particularly important when considering the formation of population structure in lotic systems. Gene flow in riverine species without a terrestrial or volant dispersal stage is restricted by the hierarchical nature of lotic networks, which offer limited dispersal paths between populations (Campbell Grant, Lowe, & Fagan, 2007; Hughes, 2007). Hence, while rates of gene flow between populations should decrease with distance, even spatially proximate populations may exhibit little to no gene flow if they belong to different catchments or basins (e.g., the Stream Hierarchy model; Meffe & Vrijenhoek, 1988). Site-level factors such as dams and impounds can further impede gene flow by fracturing habitat and isolating populations regardless of proximity within the stream network hierarchy (e.g., the Death Valley model; Finn et al., 2007; Mullen, Woods, Schwartz, Sepulveda, & Lowe, 2010).

The complexity of gene flow in lotic systems makes it difficult to disentangle the factors responsible for forming patterns of genetic variation. Thus, attempts to estimate particular evolutionary parameters using a model of evolutionary history that does not adequately account for the impacts of population divergence, gene flow and demographic history may lead to poor model fit, and inaccurate parameter estimation (Thomé & Carstens, 2016). Consequently, comprehensive approaches for studying the evolution of population structure in these systems should properly account for the role of all of these forces.

Similarly, sufficient sampling of genetic loci is a key factor in accurately estimating recent, fine-scale population genetic processes (Catchen et al., 2013; Emerson et al., 2010; Nunziata, Lance, Scott, Lemmon, & Weisrock, 2017). Studies using traditional molecular markers (e.g., mitochondrial DNA and microsatellites) have often been limited to small numbers of evolutionarily independent loci. However, more recently developed techniques, such as reduced-representation sequencing using restriction site-associated DNA (RADseq), permit the sampling of thousands of independently evolving loci without the need for prior genome sequence information (Baird et al., 2008; Hohenlohe, Amish, Catchen, Allendorf, & Luikart, 2011; Hohenlohe, Catchen, & Cresko, 2012). These techniques can thus provide the genomic information needed to tease apart processes acting to structure genetic variation at fine geographic scales (Andrews, Good, Miller, Luikart, & Hohenlohe, 2016). When coupled with analyses aimed at identifying the forces acting to structure genetic variation, these data can be highly informative about the overall evolutionary and population history of a species (Anderson et al., 2010; Catchen et al., 2013).

In this study, we investigated patterns of population structure across multiple river basins using genome-wide data generated from the common mudpuppy salamander, *Necturus maculosus* (Proteidae). While many salamander species can disperse terrestrially between streams (Miller, Snodgrass, & Gasparich, 2015), *N. maculosus* is obligately paedomorphic and wholly restricted to its aquatic environment, preventing overland dispersal (Petranka, 1998). Given this constraint, we predicted that genetic distance between populations of *N. maculosus* should increase with river distance and that population structure would be partitioned primarily by catchment and basin. However, we also tested whether idiosyncratic patterns of isolation and reduced gene flow exist as a result of extrinsic, site-specific factors such as dams. To avoid confirmation bias, we developed and tested multiple models of evolutionary history that consider divergence, gene flow and demographic history, and we examined the relative influence of spatial scale, stream network architecture and dispersal barriers (i.e., dams and impounds). Our work identifies the role of stream network hierarchy in driving population structure and also indicates how local barriers can idiosyncratically alter the connectivity between sites to generate a more heterogeneous pattern of population structure across the landscape.

2 | METHODS

2.1 | Sampling sites and design

We sampled *N. maculosus* at two hierarchical scales (river basins and sites within basins) to examine spatial patterns of population structure across eastern and central Kentucky (USA). We sampled within three major river basins (Figure 1), all of which flow directly into the Ohio River: the Licking River basin, the Kentucky River basin and the unbranched Kinniconick Creek basin located in north-eastern Kentucky. While Kinniconick Creek is not a typical high-order river, its catchment is independent of other rivers and flows directly into the Ohio River, thus occupying the same hierarchical level as the two other rivers. Hereafter, we generally refer to the Kinniconick Creek basin as a "river basin." Within each basin, we sampled from one to five sites (Table 1). The study sites varied in terms of distance to the next closest site, ranging from 6 to 268.1 km (Table S1), as measured using the National Inventory of Dams measuring tool (USACE 2013). Two sites in the Red River (a stream in the Kentucky basin) were within one km of each other and were treated as a single site for this study (Gladie). To assess for the influence of dams and impounds as potential barriers to dispersal, we chose basins with a wide range of damming, from the heavily impounded Kentucky River, to the less-disturbed Licking and Kinniconick basins. Across our study system, sites were separated by 0–13 dams (USACE 2013; Table S1).
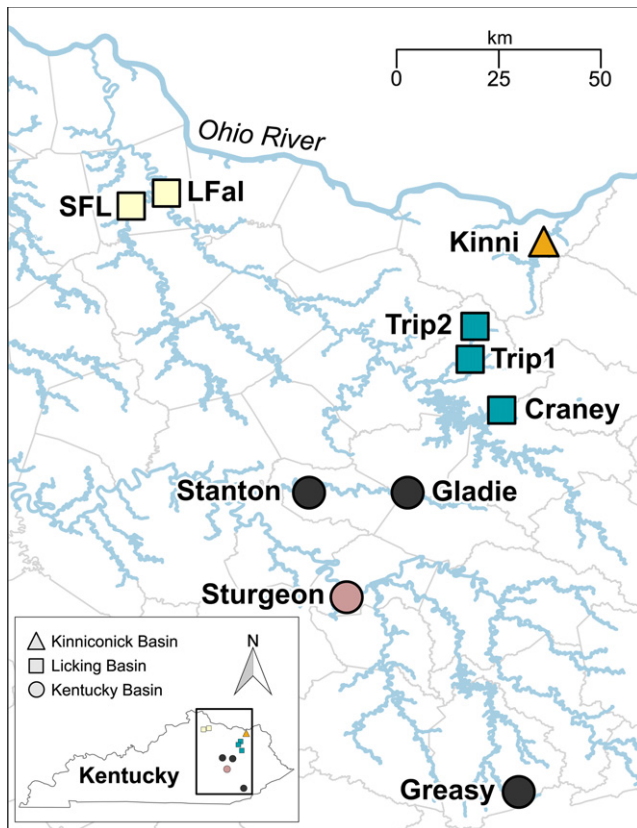
FIGURE 1 Location of sampling sites across the Kentucky, Licking and Kinniconick River basins. Shape indicates the river basin in which a sampling site is located. Colour corresponds to genetic cluster assignment from ADMIXTURE analyses under a $K = 5$ model

TABLE 1 Sites sampled with corresponding site abbreviation, site number, basin location and number of individuals captured

Site name	Site abbreviation	Basin	Number of individuals
Kinniconick Creek	Kinni	Kinniconick	5
Licking River at Falmouth	LFal	Licking	6
South Fork Licking River	SFL	Licking	4
N. Fork Triplett Creek 2	Trip2	Licking	5
Craney Creek	Craney	Licking	3
N. Fork Triplett Creek 1	Trip1	Licking	6
Gladie Creek	Gladie	Kentucky	4
Stanton Creek	Stanton	Kentucky	2
Sturgeon Creek	Sturgeon	Kentucky	5
Greasy Creek	Greasy	Kentucky	1

We captured *N. maculosus* at each site using either manual snorkel surveys or trapping, depending on the season. We collected 1–7 tissue samples per site using non-destructive tail clipping from both adult and larval individuals. Tissues were stored in 95% ethanol. A total of 41 individuals were collected from 10 sites. All tissue sampling took place between August 2013 and September 2015. For a full description of field methods, see Murphy, Price, Hime, Drayer, and Weisrock (2016). This research was approved by the Institutional Animal Care and Use Committee (protocol 2013–1073). Collections were made under Kentucky Department of Fish and Wildlife Resources Scientific Wildlife Collecting permits SC1411030 and SC1511017.

2.2 | Genetic data collection

A hurdle to the generation of population genomic data from salamanders using next generation sequencing (NGS) has been their large genome size (Gregory, 2017). This is particularly true for *N. maculosus*, which has one of the largest genomes among all vertebrates, estimated at 85 gigabases (Gb). We overcame this issue using some small adjustments and special considerations in the implementation of double-digest restriction site-associated DNA (ddRAD) sequencing (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012). This is a reduced-representation NGS method that focuses sequencing effort on a subset of the genomic fragments that are flanked on each side by restriction enzyme cutting sites. We used a larger amount of starting DNA (3,000 ng) for library preparation compared to the standard amount (100–1,000 ng) suggested in Peterson et al. (2012). Genomic DNA was digested from each individual using the restriction enzymes, SphI and EcoRI. Restriction fragments were incorporated into indexed Illumina sequencing libraries and size-selected for a mean fragment size of 376 bp (base pairs; ±10%) using a PippinPrep (Sage Science). Based on an assumption of an 85-Gb genome (Gregory, 2017), and using fragment distributions generated from both single-enzyme and double-enzyme digests, we estimated that this fragment size selection would include approximately 750,000 unique fragments from a single individual. Paired-end 150-bp sequencing was performed on pooled libraries using an Illumina HiSeq 2500. To increase the probability of recovering similar sets of orthologous loci with high read coverage across multiple individuals, we were limited to multiplexing 10–11 individuals per Illumina sequencing lane.

2.3 | ddRAD processing

We analysed our paired-end sequence reads using Stacks v.1.35 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011) to identify the total set of loci within individuals and shared orthologous loci across individuals. Given our size selection of 376 base pairs with a ±10% range, we expected no overlap between our paired-end reads; therefore, prior to processing in Stacks, we concatenated forward and reverse reads together using a custom script (script available in the Dryad accession). We filtered reads for quality using *process_radtags* in Stacks and removed reads if they contained uncalled bases, or if they contained a mean quality score < 20 within a sliding window of 15% of the read length. Reads passing quality

filtering were then de novo assembled with a minimum stack depth (i.e., number of reads) of four and a maximum of four mismatches permitted between reads. We conducted further filtering using *populations* in Stacks, by reducing the set of loci to those found in every individual (i.e., no missing data) and which had a minimum stack depth of 10 reads. To filter Stacks loci that were potentially comprised of reads from paralogous genomic loci, we used the program VCFtools v0.1.14 (Danecek et al., 2011) to remove all loci with a maximum mean read depth > 250. Post-filtered Stacks loci were used to generate both single nucleotide polymorphism (SNP) genotype matrices and multilocus sequence alignments. The former was generated using the Stacks flag –write_random_snp, which randomly selected a single SNP from each locus (when > 1 SNP was present).

2.4 | Estimating genomic diversity

We assessed genomic diversity for each study site using observed heterozygosity (H_o), expected heterozygosity (H_e) and nucleotide diversity (π). We calculated departure from random mating using Wright's inbreeding coefficient (F_{IS}) at each site. To assess for signatures of demographic expansion or contraction, we calculated Tajima's D using the PopGenome package in R v.3.2.3 (Pfeifer, Wittelsbürger, Ramos-Onsins, & Lercher, 2014). We calculated Tajima's D at the site and basin levels, and with all sites combined, using a 95% confidence interval around 0 for a rough estimate of significance (i.e., $-2 > D > 2$; Anholt & Mackay, 2009). We assessed genetic differentiation between sites using pairwise F_{ST} statistics. All summary statistics except for Tajima's D were calculated in Stacks.

2.5 | Primary assessment of population structure

We used multiple approaches to assess the broad pattern of population structure across our three-river study system. First, we used two different clustering and assignment methods that allow for the determination of the number of differentiated population genetic clusters. These methods also allow for the identification of individuals that may be admixed with genomic variation from two or more genetic clusters. For one of these clustering methods, we analysed our SNP genotype data in replicates of 10 using the model-based program ADMIXTURE (Alexander, Novembre, & Lange, 2009), which assigns individuals to genetic clusters using a maximum-likelihood approach without requiring a pre-defined assignment of individuals to populations. The best-fitting model was identified as the one with the number of clusters (K) corresponding with the lowest cross-validation error score. Results were visualised with the program Clumpak (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015).

As a second population clustering analysis, we analysed our SNP genotype data using discriminant analysis of principal components (DAPC; Jombart, Devillard, & Balloux, 2010) implemented in the adegenet package in R. This multivariate statistical approach uses a principal components analysis to transform data into a smaller set of uncorrelated variables and then partitions individuals into clusters that maximise differentiation within discriminant space. We performed 1,000 replicates of cross-validation using the R package poppr (Kamvar, Tabima, & Grünwald, 2014) to determine the number of principal components (PCs) that gave highest mean cluster assignment success and lowest mean standard error. Discriminant analyses were performed using two different levels of hierarchical structure: (1) using each of the three basins as a separate population, and (2) dividing the rivers into five populations based on high assignment in ADMIXTURE at a $K = 5$ (see Results).

We also used a phylogenetic approach to search for hierarchical patterns and identify relationships between individuals and populations across our *N. maculosus* study system. We analysed our SNP data to estimate a lineage tree for all individuals using SVDquartets implemented in PAUP* (Swofford, 2017; Chifman & Kubatko, 2014). SVDquartets infers tree topology by subsetting data into quartets, assessing the likelihood of each quartet under a coalescent model, and then combining quartets into a full tree (Chifman & Kubatko, 2014). This method reduces the amount of computation time required, compared to other commonly used phylogenetic programs, while providing comparable results (Chou et al., 2015). We were thus able to analyse our complete unlinked SNP data set to evaluate all possible quartets. In addition, we performed 1,000 bootstrap replicates to generate measures of nodal support.

We examined the partitioning of genetic variation across different hierarchical scales of population structure using an analysis of molecular variance (AMOVA; Excoffier, Smouse, & Quattro, 1992) implemented in the R package pegas (Meirmans, 2006; Paradis, 2010). As with DAPC, we performed two different analyses: (1) an analysis that identified the uppermost level of hierarchical structure using river basins, and (2) an analysis where this was defined based on results from the above-described population structure analyses. In all analyses, we calculated the degree to which genetic variation was partitioned across three levels: between basins or genetic clusters, among sites within basins or within genetic clusters, and among individuals within localities.

2.6 | Testing for a history of lineage divergence

Although *N. maculosus* is considered a single species, it is important to consider the potential for the historical forces of lineage divergence in generating patterns of population structure, particularly given the history of the ancestral Teays River and its impact on the formation of the contemporary Ohio River network during late Pleistocene Glaciation (Berendzen, Simons, & Wood, 2003; Teller, 1973). We used the program Bayesian Phylogenetics and Phylogeography (BPP) v3.3 to test hypotheses of lineage divergence (Yang & Rannala, 2010). BPP performs unguided species delimitation, jointly estimating both the species tree and species delimitation under a multispecies coalescent model in a Bayesian framework. BPP requires the identification of the maximum number of lineages that could be present in a system; branches cannot be split, but the program allows for collapsing of nodes to combine branches into more inclusive lineages, and allows for the regrafting of branches to alter topology (Yang, 2015). Thus, we used the most liberal hypothesis of lineage

divergence using results from ADMIXTURE and assigning each cluster recovered from ADMIXTURE to its own lineage. The number of individuals was equalised across branches by randomly selecting five individuals from each cluster. We used a guide tree topology that was extrapolated from the lineage tree created in SVDquartets. While the topology of the guide tree should theoretically have little influence on the results, we eliminated the possibility for topology biases by running analyses twice: first with the SVDquartet tree, and then with a randomised tree (Carstens, Pelletier, Reid, & Satler, 2013; Yang & Rannala, 2014). BPP uses multilocus sequence alignments as input data. To overcome computation restrictions, we used a randomly selected subset of 1,877 loci (approximately half of all loci) for all BPP analyses. Reverse-jump Markov chain Monte Carlo (MCMC) analyses were run for a total of 1,000,000 generations, after a burn-in stage of 10,000 generations. Topology and parameter estimates were sampled every 10 generations for a total of 100,000 samples. To ensure that the divergence and population size priors were not biasing results, we performed these analyses using a range of different priors under each of our starting trees. To assess convergence, four replicate analyses were performed for each combination of starting tree and priors, using the same subset of loci for each replicate.

2.7 | Testing for isolation by distance and isolation by barriers

We tested for a signature of isolation by distance (IBD), the effects of damming, and for correlations between river distance and damming using Mantel tests of correlations between genetic distance and river distance, genetic distance and number of dams between sites, and river distance and number of dams between sites, respectively. We also performed partial Mantel tests (Smouse, Long, & Sokal, 1986) to test for a correlation between genetic distance and number of dams between sites while controlling for either geographic distance or basin assignment, and to test for a correlation between genetic distance and geographic distance while controlling for either the number of dams between sites or river basin. All Mantel and partial Mantel tests were performed in the R *vegan* package (Oksanen et al., 2007) using 9,999 matrix randomisations.

2.8 | Demographic model testing and gene flow estimation

We performed demographic model testing and parameter estimation using a coalescent-based Bayesian MCMC approach in Migrate-n (Beerli & Palczewski, 2010). Migrate-n allows for the assessment of models that account for the parameters of effective population size (Θ) and gene flow (M) and their effect on genetic variation within and between populations. Marginal likelihoods calculated for models that vary in gene flow between populations (no migration, unidirectional migration or bidirectional migration) can be used to identify the best-fitting model for a set of populations, and parameter estimates can then be drawn from the best-fitting model.

We focused this work on eight models that explored gene flow between the three river basins (Figure S1), and as such, treated each river basin as a population. Models that account for gene flow among sites within basins were too parameter rich to be completed in a feasible amount of time. Furthermore, due to computational limitations, analyses were limited to a randomly selected subset of 1,000 loci. Loci were analysed as sequence alignments, and not as SNPs. For each set of models, test runs were performed using the full migration model to determine an appropriate range of priors and these values were then standardised across all models. All subsequent runs were performed using four Markov chains and a static heating scheme (1.0, 1.5, 3.0 and 1,000,000) swapping every 10 generations. Chains were run for 10,000,000 generations, with the first 1,000,000 generations discarded as burn-in, and with samples drawn every 100 generations for a total of 10,000 recorded steps. Ten replicates were performed for each model to ensure that results were consistent between runs and provide evidence of convergence. We calculated Bayes factors (BF) from Bezier log marginal likelihoods as two times the difference in marginal likelihoods between the best-fitting model and alternate models. Bayes factors >10 were considered as decisive support in favour of the best-fitting model (Kass & Raftery, 1995).

3 | RESULTS

3.1 | Genetic diversity

A total of 1,439,623 loci were identified across all 41 individuals. Quality filtering resulted in 6,873 loci, with an average length of 290 bp, shared across all individuals with no missing data. In all sampling localities, H_o was higher than H_e (Table 2); however, the variance on these calculations was high and there was no significant difference between the two. Measures of π were generally similar across localities. F_{IS} values were generally negative, but not significantly different from zero (Table 2). Pairwise F_{ST} between sampling sites ranged from 0.061 to 0.264 and were generally greater between river basins (Table 3). When calculated across all populations, Tajima's D was -0.78. Tajima's D at the basin and site levels were also generally negative, although a few sites had positive values (Table S2). Overall, no value of Tajima's D was substantially large or small (i.e., < -2 or > 2). Private alleles were roughly similar across all sampled sites.

3.2 | Population structure

ADMIXTURE results best supported the assignment of individuals into $K = 2$ genetic clusters based on cross-validation error scores (Figure S2); however, models for $K = 3$–5 were only slightly less supported based on this metric. The $K = 2$ model separated the Kentucky basin from the combined Kinniconick and Licking basins (Figure 2). The $K = 3$ model further separated the Sturgeon Creek site from the other Kentucky basin sites. The $K = 4$ model further separated the Kinniconick basin from the Licking basin. Finally, the

TABLE 2 Population genetic summary statistics for all 10 sampling sites, including observed heterozygosity (H_o), expected heterozygosity (H_e), nucleotide diversity (π). Site abbreviations are as in Table 1

Site	H_o	H_e	π	F_{IS}	Private alleles
Kinniconick	0.120	0.100	0.111	−0.0201	512
LFal	0.137	0.121	0.132	−0.0104	582
SFL	0.119	0.098	0.112	−0.0131	363
Trip2	0.116	0.102	0.114	−0.0039	199
Craney	0.097	0.091	0.109	0.0202	239
Trip1	0.113	0.094	0.103	−0.0233	167
Gladie	0.147	0.126	0.144	−0.0063	577
Stanton	0.133	0.109	0.145	0.0192	409
Sturgeon	0.121	0.095	0.106	−0.032	415
Greasy	0.111	0.055	0.111	0	163

$K = 5$ model further separated the combined SFL and LFal sites from the remaining Licking basin sites. Overall, individuals were assigned to their clusters with high probability across all levels of K from 2–5 (Figure 2) with a small number of individuals having assignments indicating low levels of admixture between clusters (e.g., individuals sampled from the LFal site in the Licking basin).

Discriminant analysis of principal components cross-validation resulted in 19 PCs and 2 discriminant axes (DAs) being retained for the three-population analysis, and 20 PCs and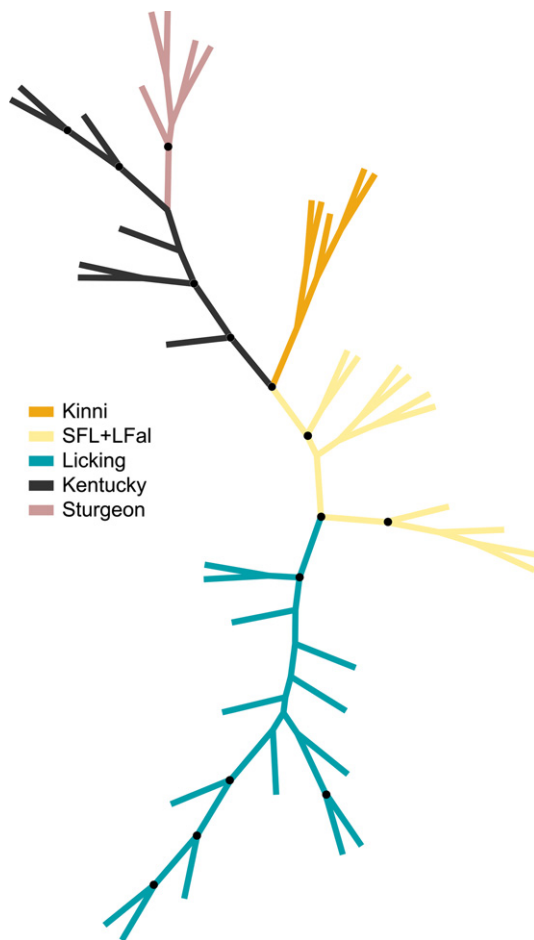 4 DAs being retained for the five-population analysis. Results from the three-population analysis showed strong differentiation on axis 1 between the Kentucky basin and the other two river basins (Figure 3a), and further differentiation between the Kinniconick and Licking basins on axis 2. Results from the five-population analysis identified the greatest amount of differentiation on axis 1 between the Sturgeon Creek site within the Kentucky basin and the other genetic clusters (Figure 3b). Additional genetic differentiation existed among clusters on axis 2, but to a much less extent than that seen on axis 1. While most individuals were assigned to the cluster primarily associated with their river basin, some posterior assignments showed discordance between genetics and geography, with individuals being assigned to a genetic cluster different than the geographic population from

where they were sampled. For example, in both the three- and five-population analyses one individual sampled from the Kentucky basin was given a strong membership probability to a genetic cluster primarily assigned to Licking basin individuals.

Exhaustive quartet sampling in SVDquartets produced an unrooted lineage tree that clustered populations into different regions of the tree, but without the formation of monophyletic groups corresponding to all genetic clusters identified in Admixture and DAPC analyses (Figure 4). River basins did form monophyletic groups, with the Kentucky basin and Licking basin both receiving strong bootstrap support (> 95%).

In all AMOVAs, the majority of genetic variation was explained by variation among individuals within localities (> 83%; Table 4). Among-basin variation was not significant when treating the three individual river basins as the uppermost level of population structure ($p = .145$). It was significant at the among-basin level when the Licking and Kinniconick basins were combined into a single group ($p = .018$); however, this explained a relatively small percentage of the overall genetic variation (6.7%).

3.3 | Tests of lineage divergence

Coalescent-based tests of lineage divergence rejected all hypotheses of lineage divergence within our study system. A model placing all populations into a single lineage was consistently supported with a posterior probability of 1.0 in all analyses, including those using different starting guide tree topologies and prior settings. All resulting output files from these analyses can be found in the Dryad accession.

3.4 | IBD and isolation by barriers

Mantel tests applied to the total study system resulted in a positive correlation between genetic distance and geographic distance ($p < .001$, $r = .607$), and a positive correlation between genetic distance and number of dams between sites ($p < .001$, $r = .715$). Partial Mantel tests resulted in a significant correlation between genetic distance and dams when controlling for geographic distance ($p < .019$, $r = .484$). Partial Mantel tests also found significant correlations between geographic and genetic distances when controlling for river basin ($p = .044$, $r = .404$), and between genetic distance and the

TABLE 3 Pairwise F_{ST} for all 10 sampling sites

	LFal	SFL	Trip2	Craney	Trip1	Gladie	Stanton	Greasy	Sturgeon
Kinni	0.078	0.121	0.104	0.121	0.118	0.121	0.161	0.209	0.177
LFal		0.068	0.061	0.071	0.071	0.087	0.112	0.139	0.133
SFL			0.104	0.128	0.114	0.120	0.166	0.226	0.181
Trip2				0.082	0.065	0.111	0.149	0.196	0.167
Craney					0.094	0.124	0.182	0.264	0.192
Trip1						0.123	0.164	0.212	0.181
Gladie							0.099	0.139	0.124
Stanton								0.240	0.164
Greasy									0.211

FIGURE 2 Assignment plots resulting from ADMIXTURE analyses for models of K = 2–5. Vertical bars represent individuals with the proportion of their colour representing the level of probability of assignment to a particular genetic cluster. Abbreviations at the bottom of the plots refer to sample localities listed in Table 1
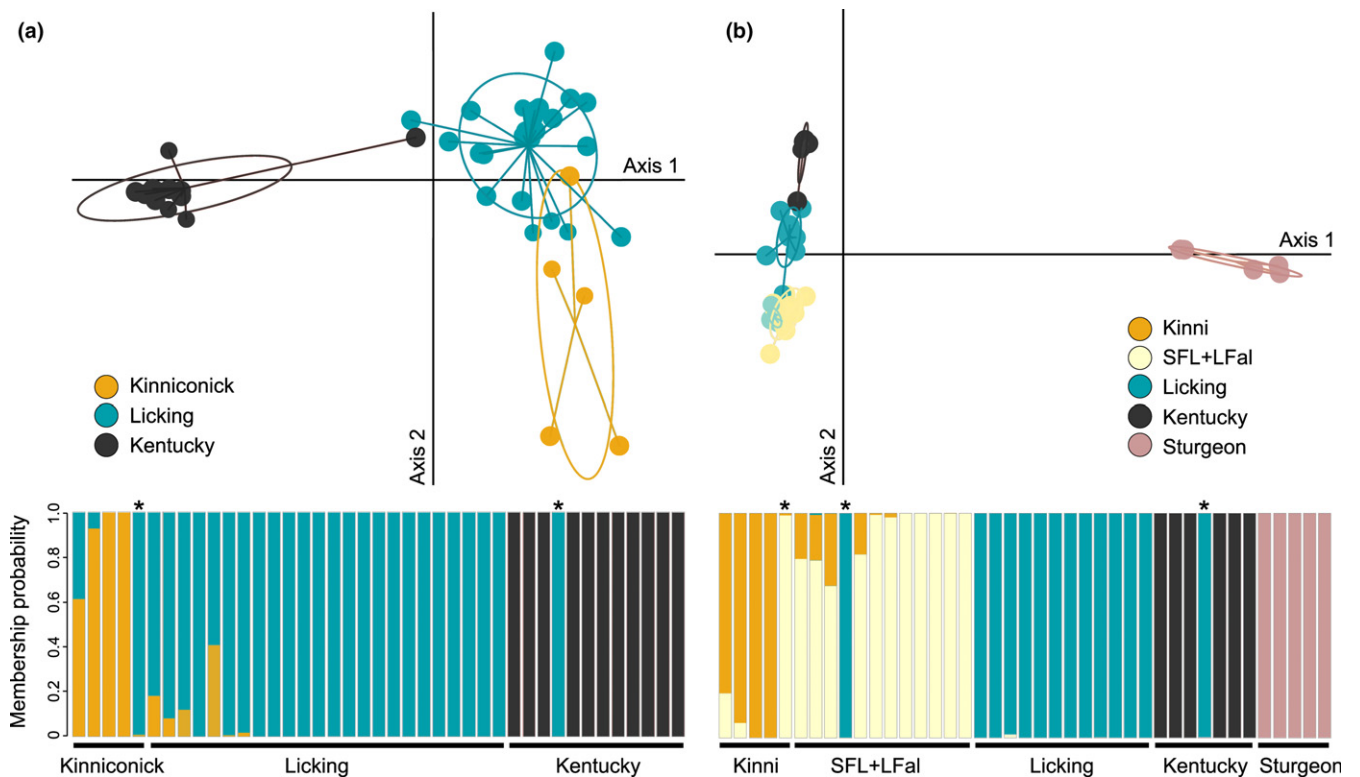
FIGURE 3 DAPC scatterplot for two clusters (a) and five clusters (b), with assignment plots indicating probability of assignment to a particular cluster. Asterisks indicate assignment of a site to a different cluster than the geographic population from where they were sampled

distance (p = .17, r = .453) or between genetic distance and the number of dams (p = .17, r = .490).

3.5 | Demographic model testing and gene flow estimation

Migrate-n results produced decisive support for a model that included bidirectional gene flow between the Kentucky and Licking basins, and no gene flow between the Kinniconick basin and either the Kentucky or Licking basins (Figure 5). This model was favoured over the next best model, which allowed for bidirectional migration between all rivers, with a BF = 4418, indicating a substantially worse fit to our data (Table 5). Under the best-fit model, the Kentucky basin had the largest effective population size with Θ = 0.0044 (95% CI: 0.0027, 0.0060), followed by the Licking River with Θ = 0.0023 (0.00067, 0.0040) and the Kinniconick with Θ = 0.00021. The mean migration rate (M) was 764 (400, 1,120) from the Kentucky basin into the Licking basin and 7,501 (7,000, 7,960) from the Licking to the Kentucky. We note that all estimates of Θ and M are scaled by the mutation rate. Conversion of these two parameters into an estimate of the mean number of migrants per generation (Nm) produced an Nm = 0.45 from the Kentucky basin into the Licking basin, and an Nm = 8.16 from the Licking basin into the Kentucky basin.

4 | DISCUSSION

Within our *N. maculosus* study system, we used a genomic approach to test the prediction that population structure should be positively correlated with river distance and that some populations may be disproportionately affected by local factors. Our results demonstrated that population structure is indeed influenced by river network architecture, but that substantial patterns of isolation and limited gene flow also exist that are not explained by hierarchical network architecture. Consequently, even on this relatively fine scale, population structure cannot be explained by a single model.

4.1 | Tests of lineage divergence

The quantification of gene flow and demography can be strongly biased if historical lineage divergence was a major factor in the

FIGURE 4 SVDquartets unrooted lineage tree of all individuals sampled. Branches are coloured based on population assignments from ADMIXTURE. Nodes marked with filled black circles have > 95% bootstrap support

Legend:
- Kinni
- SFL+LFal
- Licking
- Kentucky
- Sturgeon

number of dams when controlling for the effect of basin (p = .009, r = .7). No significant correlation was found between genetic distance and geographic distance when controlling for the effect of dams between sites (p = .692, r = −.105).

Partial Mantel tests applied to the combined Licking and Kinniconick basins found a positive correlation between genetic and geographic distances when controlling for the number of dams (p < .044, r = .483). Within the Kentucky basin, no significant relationship was detected between genetic distance and geographic

TABLE 4 Results from analysis of molecular variance using hierarchical structure according to river basin and results from population structure analyses

Level of structure	Source of Variation	Variance components	Percentage of variation	Fixation indices	p
Kentucky	Among rivers	0.854	2.953	0.03	.145
versus Licking	Among localities within rivers	6.033	11.799	0.122	.001
versus Kinniconick	Among samples within localities	−9.316	85.248	0.148	.001
Kentucky	Among clusters	3.364	6.745	0.067	.018
versus Licking	Among sites within clusters	6.095	10.157	0.109	.001
+ Kinniconick	Among samples within sites	−9.475	83.098	0.169	.001

TABLE 5 Demographic model selection for analyses of three-population models treating the Kentucky, Kinnicionick and Licking basins as separate populations. Models are listed according to their level of support from marginal likelihoods and Bayes factors. See Figure S1 in the Supporting Information for the full model schematic with corresponding model labels

Model	Bezier approximation marginal likelihood	Bayes factor	Interpretation
G	−686208	0	No migration TO/FROM Kinni
A	−690626	−4418	FULL MODEL
D	−690910	−4703	No migration between Kentucky and Kinni
B	−691051	−4843	No migration between Kentucky and Licking
C	−692022	−5814	No migration between Licking and Kinni
E	−963610	−277402	No migration TO/FROM Licking
H	−1145516	−459308	No migration
F	−1152748	−466540	No migration TO/FROM Kentucky
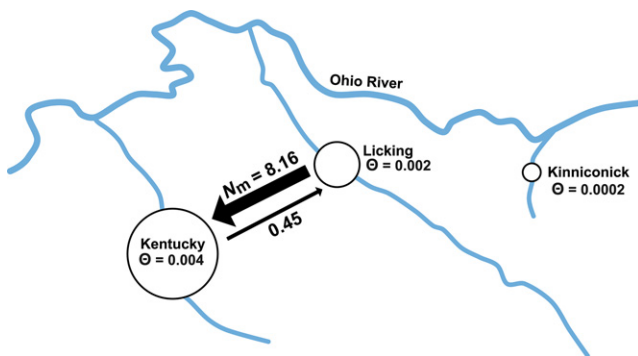
FIGURE 5 Schematic of the best-fitting Migrate-n model showing estimated effective population sizes (Θ) of each basin and the mean number of migrants per generation between basins (*Nm*)

structuring of genetic variation (Thomé & Carstens, 2016). Hence, when clear geographic patterns of genetic structure are identified, it is important to determine whether these patterns reflect historical lineage divergence or are evidence of more recent population structure. It is possible that historical lineage divergence could have played a role in our *N. maculosus* study system. During the Pleistocene, the Kentucky basin was connected with the old Ohio River basin, and the Licking and Kinnicionick basins were connected to the north-flowing Teays River palaeodrainage (Teller, 1973; Teller & Goldthwait, 1991). These palaeodrainage patterns have been linked to the population structure and divergence of a wide range of aquatic species, including fish, aquatic insects and salamanders (Berendzen et al., 2003; Kozak, Blaine, & Larson, 2006; Kuchta,

Haughey, Wynn, Jacobs, & Highton, 2016; Pessino, Chabot, Giordano, & DeWalt, 2014).

Despite this historical potential, we found no evidence for a history of lineage divergence. Coalescent-based tests strongly rejected models of lineage divergence among any of the individual or combined genetic clusters identified in the exploratory analyses of population structure, giving us confidence in treating our study system as a set of populations linked together by restricted gene flow. These results are congruent with a study of mtDNA variation in *N. maculosus* (Stedman, 2016), which identified a very shallow phylogeographic history across its eastern range, with no signature of divergence that could be linked to the historical Teays palaeodrainage.

4.2 | Population structure and gene flow in *N. maculosus*

In general, population structure of *N. maculosus* in our study system appears to be influenced by network architecture, with genetic differentiation increasing with river distance, and with some of the strongest patterns of differentiation occurring between river basins. This pattern was evident in our explorations of population structure and general summary statistics (i.e., F_{ST}), as well as results from simple Mantel tests, which indicated a correlation between geographic and genetic distance across the study system. Species restricted to a continuously connected dendritic network generally are expected to follow this pattern [e.g., the Stream Hierarchy Model (Meffe & Vrijenhoek, 1988)], and population genetic studies in fully aquatic salamander species have consistently followed this expectation, including investigations of Cope's giant salamander, *Dicamptodon copei* (Dicamptodontidae; Steele, Baumsteiger, & Storfer, 2009), and the hellbender, *Cryptobranchus alleganiensis* (Cryptobranchidae; Unger, Rhodes, Sutton, & Williams, 2013). *Necturus maculosus* may have the ability to disperse long distances; it can inhabit a variety of habitats (Petranka, 1998) and is suggested to exhibit seasonal migration (Matson, 1998). However, it has also been shown to exhibit synchronised upstream movements that may be related to spawning (Green & Pauley, 1987; Pfingsten et al., 2013; Shoop & Gunning, 1967). This, coupled with high nest site fidelity (Matson, 1998; Shoop & Gunning, 1967), may constrain dispersal and yield the strong differentiation we see across basins.

However, not all of our results point to an exclusive link between genetic and riverine distance. Indeed, numerous lines of evidence identify patterns of population structure that are not influenced by river network architecture. At the broadest scale, while the three river basins are each identified as genetically differentiated groups in our exploratory analyses of population structure, AMOVAs indicated that differentiation among the three river basins only accounted for ∼3% of the total genetic variation. Differentiation within river basins accounted for almost 12% of genetic variation, indicating a substantial effect for the reduction in gene flow within river basins. In addition, across the total study system, we found a significant effect for the role of dams in driving genetic

differentiation, even after controlling for the effect of geographic distance. Furthermore, no significant correlation was found between genetic and geographic distances after controlling for the effect of dams. Dams significantly affect the dispersal and population structure of aquatic species (Fullerton et al., 2010; Neraas & Spruell, 2001; Nislow, Hudy, Letcher, & Smith, 2011), and populations of aquatic species in impounded rivers have been shown to have significantly increased population genetic structure relative to unimpounded rivers (Bessert & Ortí, 2008). Although the effects of damming have been widely reported in fish species, our study suggests that the damming could possibly affect obligately aquatic salamanders in a similar manner.

One of the more interesting patterns that surfaced in our results was the great genetic distinctiveness of the Sturgeon Creek population within the Kentucky River basin. DAPC results based on the assignment of individuals to five genetic clusters were particularly surprising in separating Sturgeon Creek from all other genetic groups to a greater extent than that between any other group (Figure 3b). The reasons underlying this strong pattern of differentiation for Sturgeon Creek are unclear. Other populations within the Kentucky River have as large or larger F_{ST} values when compared to sites within or outside of this basin. Measures of genetic diversity for Sturgeon Creek, such as heterozygosity or nucleotide diversity, do not seem to differentiate it from other sampling localities, indicating that recent demographic changes (e.g., population decline) are not the cause of its genetic distinctiveness. Further work will be required to better understand this result, including sampling additional *N. maculosus* localities within Sturgeon Creek.

In our model-based analysis of gene flow treating the three basins as separate populations, the top supported model allowed for bidirectional gene flow between the Kentucky and Licking basins, but contained no gene flow between the Kinniconick basin and either the Kentucky or Licking basins. The connection between the Kentucky and Licking basins is consistent with river flow, with substantially higher rates of gene flow from the Licking basin into the Kentucky basin ($Nm = 8.16$). This is higher than the standard benchmark minimum of $Nm = 1$ generally considered sufficient to prevent differentiation between populations; however, violations of the many assumptions of the one migrant per generation rule can result in higher Nm levels, while still resulting in population differentiation (Mills & Allendorf, 1996).

The lack of model-based support for gene flow in or out of the Kinniconick River basin was surprising given its close association with the Licking River basin in the ADMIXTURE and DAPC results. However, the Kinniconick basin was distinct when exploring higher levels of genetic clustering (i.e., $\geq K = 4$), which indicated a partitioning of genetic variation at a much finer level. We suggest that these results are consistent with the Kinniconick basin becoming relatively recently isolated from the other river basins in our study. We attempted to perform more focused model-based examinations of gene flow between the Licking and Kinniconick basins, which could have been useful in further clarifying connectivity among proximal basins. However, these analyses exhibited very poor signs of

convergence on the posterior distribution and are not presented. Future work with greater sampling, potentially including individuals sampled from the Ohio River, may be useful in better informing the fine-scale patterns of gene flow between basins.

5 | CONCLUSIONS

To effectively conserve riverine landscapes and species, we must first understand how landscape features and life history traits have worked in concert to shape the current distribution and population structure of species. This requires teasing apart patterns of population structure and gene flow at multiple evolutionary and spatial scales. The comparative study of other freshwater taxa across our study system, such as fish and freshwater mussels, will be useful in identifying whether the patterns of population structure identified here, especially the isolation of the Sturgeon Creek site, are unique to *N. maculosus*, or shared across aquatic communities. *Necturus maculosus* serves as a host for an imperiled unionid mussel (*Simpsonaias ambigua*), the only known mussel-host system that involves a non-fish host, and an understanding of population structure and gene flow of both species could be particularly informative of the shared factors affecting these species (Zanatta & Wilson, 2011).

AUTHOR CONTRIBUTIONS

M.M., S.P. and D.W. conceived the project. M.M. Collected the samples, conducted genomic library preparation, conducted data analysis and led manuscript preparation. K.J. conducted data analysis, including population structure and gene flow analyses. S.P. and D.W. supervised data analysis and interpretation. All authors contributed to writing the final manuscript.

DATA ACCESSIBILITY

ADMIXTURE, DAPC, BPP, SVDquartets input files, sequence stitch script and SVDquartets output files: Dryad doi: https://doi.org/10.5061/dryad.hj093

ORCID

Mason O. Murphy (iD) http://orcid.org/0000-0002-5317-1386

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Anderson, C. D., Epperson, B. K., Fortin, M. J., Holderegger, R., James, P., Rosenberg, M. S., ... Spear, S. (2010). Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, *19*, 3565–3575. https://doi.org/10.1111/j.1365-294X.2010.04757.x

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*, 81–92. https://doi.org/10.1038/nrg.2015.28

Anholt, R. R., & Mackay, T. F. (2009). *Principles of behavioral genetics*. Oxford, UK: Academic Press.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., ... Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One*, *3*, e3376.

Beerli, P., & Palczewski, M. (2010). Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*, *185*, 313–326.

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

Bessert, M. L., & Ortí, G. (2008). Genetic effects of habitat fragmentation on blue sucker populations in the upper Missouri River (*Cycleptus elongatus* Lesueur, 1918). *Conservation Genetics*, *9*, 821–832. https://doi.org/10.1007/s10592-007-9401-4

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*, 165–175. https://doi.org/10.1111/j.1461-0248.2006.01007.x

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Molecular Ecology*, *22*, 4369–4383. https://doi.org/10.1111/mec.12413

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182. https://doi.org/10.1534/g3.111.000240

Catchen, J., Bassham, S., Wilson, T., Currey, M., O'Brien, C., Yeates, Q., & Cresko, W. A. (2013). The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Molecular Ecology*, *22*, 2864–2883. https://doi.org/10.1111/mec.12330

Chifman, J., & Kubatko, L. (2014). Quartet inference from SNP data under the coalescent model. *Bioinformatics*, *30*, 3317–3324. https://doi.org/10.1093/bioinformatics/btu530

Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., & Warnow, T. (2015). A comparative study of SVDquartets and other

coalescent-based species tree estimation methods. *BMC Genomics*, *16*, S2. https://doi.org/10.1186/1471-2164-16-S10-S2

Coulon, A., Guillot, G., Cosson, J. F., Angibault, J. M. A., Aulagnier, S., Cargnelutti, B., ... Hewison, A. J. M. (2006). Genetic structure is influenced by landscape features: Empirical evidence from a roe deer population. *Molecular Ecology*, *15*, 1669–1679. https://doi.org/10.1111/j.1365-294X.2006.02861.x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Emerson, K. J., Merz, C. R., Catchen, J. M., Hohenlohe, P. A., Cresko, W. A., Bradshaw, W. E., & Holzapfel, C. M. (2010). Resolving post-glacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 16196–16200. https://doi.org/10.1073/pnas.1006538107

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*, 479–491.

Finn, D. S., Blouin, M. S., & Lytle, D. A. (2007). Population genetic structure reveals terrestrial affinities for a headwater stream insect. *Freshwater Biology*, *52*, 1881–1897.

Fullerton, A. H., Burnett, K. M., Steel, E. A., Flitcroft, R. L., Pess, G. R., Feist, B. E., ... Sanderson, B. L. (2010). Hydrological connectivity for riverine fish: Measurement challenges and research opportunities. *Freshwater Biology*, *55*, 2215–2237.

Green, N. B., & Pauley, T. K. (1987). *Amphibians and reptiles in West Virginia*. Pittsburgh, PA: University of Pittsburgh Press.

Gregory, T. R. (2017). *Animal genome size database*. Retrieved from http://www.genomesize.com.

Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., & Luikart, G. (2011). Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources*, *11*, 117–122. https://doi.org/10.1111/j.1755-0998.2010.02967.x

Hohenlohe, P. A., Catchen, J., & Cresko, W. A. (2012). Population genomic analysis of model and nonmodel organisms using sequenced RAD tags. In: F. Pompanon & A. Bonin (Eds.), *Data production and analysis in population genomics. Methods in molecular biology (methods and protocols)* (Vol. *888*, pp. 235–260). Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-61779-870-2

Hughes, J. M. (2007). Constraints on recovery: Using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology*, *52*, 616–631. https://doi.org/10.1111/j.1365-2427.2006.01722.x

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*, 1898–1914. https://doi.org/10.1111/j.1558-5646.1999.tb04571.x

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94. https://doi.org/10.1186/1471-2156-11-94

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/peerj.281

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes

and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*, 1179–1191. https://doi.org/10.1111/1755-0998.12387

Kozak, K. H., Blaine, R. A., & Larson, A. (2006). Gene lineages and eastern North American palaeodrainage basins: Phylogeography and speciation in salamanders of the Eurycea bislineata species complex. *Molecular Ecology*, *15*, 191–207.

Kuchta, S. R., Haughey, M., Wynn, A. H., Jacobs, J. F., & Highton, R. (2016). Ancient river systems and phylogeographical structure in the spring salamander, *Gyrinophilus porphyriticus*. *Journal of Biogeography*, *43*, 639–652. https://doi.org/10.1111/jbi.12668

Matson, T.O. (1998). Evidence for home ranges in mudpuppies and implications for impacts due to episodic applications of the lampricide TFM. In M.J. Lannoo (Ed.), *Status and conservation of Midwestern amphibians* (pp. 278–287). Iowa City, IA: University of Iowa Press.

Meffe, G. K., & Vrijenhoek, R. C. (1988). Conservation genetics in the management of desert fishes. *Conservation Biology*, *2*, 157–169. https://doi.org/10.1111/j.1523-1739.1988.tb00167.x

Meirmans, P. G. (2006). Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, *60*, 2399–2402. https://doi.org/10.1111/j.0014-3820.2006.tb01874.x

Miller, W. L., Snodgrass, J. W., & Gasparich, G. E. (2015). The importance of terrestrial dispersal for connectivity among headwater salamander populations. *Ecosphere*, *6*, 1–9.

Mills, L. S., & Allendorf, F. W. (1996). The one-migrant-per-generation rule in conservation and management. *Conservation Biology*, *10*, 1509–1518. https://doi.org/10.1046/j.1523-1739.1996.10061509.x

Mullen, L. B., Woods, H. A., Schwartz, M. K., Sepulveda, A. J., & Lowe, W. H. (2010). Scale-dependent genetic structure of the Idaho giant salamander (*Dicamptodon aterrimus*) in stream networks. *Molecular Ecology*, *19*, 898–909. https://doi.org/10.1111/j.1365-294X.2010.04541.x

Murphy, M. O., Price, S. J., Hime, P. A., Drayer, A. N., & Weisrock, D. W. (2016). A review of common mudpuppy (*Necturus maculosus*) capture methods and a description of a revised trap design. *Herpetological Review*, *47*, 575–578.

Neraas, L. P., & Spruell, P. (2001). Fragmentation of riverine systems: The genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. *Molecular Ecology*, *10*, 1153–1164. https://doi.org/10.1046/j.1365-294X.2001.01269.x

Nislow, K. H., Hudy, M., Letcher, B. H., & Smith, E. P. (2011). Variation in local abundance and species richness of stream fishes in relation to dispersal barriers: Implications for management and conservation. *Freshwater Biology*, *56*, 2135–2144. https://doi.org/10.1111/j.1365-2427.2011.02634.x

Nunziata, S. O., Lance, S. L., Scott, D. E., Lemmon, E. M., & Weisrock, D. W. (2017). Genomic data detect corresponding signatures of population size change on an ecological time scale in two salamander species. *Molecular Ecology*, *26*, 1060–1074. https://doi.org/10.1111/mec.13988

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. A. S. S. (2007). The vegan package. *Community Ecology Package*, *10*, 631–637.

Paradis, E. (2010). pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*, 419–420. https://doi.org/10.1093/bioinformatics/btp696

Pessino, M., Chabot, E. T., Giordano, R., & DeWalt, R. E. (2014). Refugia and postglacial expansion of *Acroneuria frisoni* Stark & Brown (plecoptera: perlidae) in North America. *Freshwater Science*, *33*, 232–249. https://doi.org/10.1086/675306

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*, e37135. https://doi.org/10.1371/journal.pone.0037135

Petranka, J. W. (1998). *Salamanders of the United States and Canada*. Washington, DC: Smithsonian Institution Press.

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, *31*, 1929–1936. https://doi.org/10.1093/molbev/msu136

Pfingsten, R. A., Davis, J. G., Matson, T. O., Lipps, G. J., Wynn, D. E., & Armitage, B. J. (Eds). (2013). *Amphibians of Ohio*. Columbus, OH: Ohio Biological Survey.

Shoop, C. R., & Gunning, G. E. (1967). Seasonal activity and movements of *Necturus* in Louisiana. *Copeia*, *1967*, 732–737. https://doi.org/10.2307/1441882

Smouse, P. E., Long, J. C., & Sokal, R. R. (1986). Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, *35*, 627–632. https://doi.org/10.2307/2413122

Stedman, A. L. (2016). Phylogeographic and environmental DNA analysis of the mudpuppy (*Necturus maculosus maculosus*). Master's thesis, Eastern Michigan University.

Steele, C. A., Baumsteiger, J., & Storfer, A. (2009). Influence of life-history variation on the genetic structure of two sympatric salamander taxa. *Molecular Ecology*, *18*, 1629–1639. https://doi.org/10.1111/j.1365-294X.2009.04135.x

Swofford, D. L. (2017). PAUP* Version 4.0a159. Phylogenetic Analysis Using Parsimony (*and Other Methods). Distributed by the author. Retrieved from http://paup.phylosolutions.com

Teller, J. T. (1973). Preglacial (Teays) and early glacial drainage in the Cincinnati area, Ohio, Kentucky, and Indiana. *Geological Society of America Bulletin*, *84*, 3677–3688. https://doi.org/10.1130/0016-7606(1973)84<3677:ptaegd>2.0.co;2

Teller, J. T., & Goldthwait, R. P. (1991). The old Kentucky River; a major tributary to the Teays River. *Geological Society of America Special Papers*, *258*, 29–42. https://doi.org/10.1130/SPE258

Thomé, M. T. C., & Carstens, B. C. (2016). Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 8010–8017. https://doi.org/10.1073/pnas.1601064113

Unger, S. D., Rhodes, O. E. Jr, Sutton, T. M., & Williams, R. N. (2013). Population genetics of the eastern hellbender (*Cryptobranchus alleganiensis alleganiensis*) across multiple spatial scales. *PLoS ONE*, *8*, e74180. https://doi.org/10.1371/journal.pone.0074180

USACE. US Army Corps of Engineers. (2013). *CorpsMap: National inventory of dams*. Retrieved from http://nid.usace.army.mil/cm_apex/f?p=838:1:0::NO. Accessed 21 July 2014.

Wright, S. (1949). The genetical structure of populations. *Annals of Eugenics*, *15*, 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*, 854–865. https://doi.org/10.1093/czoolo/61.5.854

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 9264–9269. https://doi.org/10.1073/pnas.0913022107

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*, 3125–3135. https://doi.org/10.1093/molbev/msu279

Zanatta, D. T., & Wilson, C. C. (2011). Testing congruency of geographic and genetic population structure for a freshwater mussel (Bivalvia: Unionoida) and its host fish. *Biological Journal of the Linnean Society*, *102*, 669–685. https://doi.org/10.1111/j.1095-8312.2010.01596.x

Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: A review. *Landscape Ecology*, *27*, 777–797. https://doi.org/10.1007/s10980-012-9737-0

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.