

# Discerning structure versus speciation in phylogeographic analysis of Seepage Salamanders (*Desmognathus aeneus*) using demography, environment, geography, and phenotype

Alex Pyron<sup>1</sup>, Anvith Kakker<sup>2</sup>, David Beamer<sup>3</sup>, and Kyle O'Connell<sup>4</sup>

<sup>1</sup>The George Washington University

<sup>2</sup>Thomas Jefferson High School for Science and Technology

<sup>3</sup>East Carolina University

<sup>4</sup>Deloitte LLP

August 7, 2023

## Abstract

Numerous mechanisms drive ecological speciation, including isolation by adaptation, barrier, distance, environment, hierarchy, and resistance. These promote genetic and phenotypic differentiation of local populations, formation of phylogeographic lineages, and ultimately, completed speciation via reinforcement. In contrast, it is possible that similar mechanisms might lead to lineage cohesion through stabilizing rather than diversifying ecomorphological selection and the long-term persistence of population structure within species. Processes that drive the formation and maintenance of geographic genetic diversity while facilitating high rates of migration and limiting phenotypic divergence may thereby result in population structure that is not accompanied by divergence towards reproductive isolation. We suggest that this framework can be applied more broadly to address the classic dilemma of “structure versus speciation” when evaluating phylogeographic diversity, unifying population genetics, species delimitation, and the underlying study of speciation. We demonstrate one such instance in the Seepage Salamander (*Desmognathus aeneus*) from the southeastern United States. Recent studies estimated up to 6.3% mitochondrial divergence and 4 phylogenomic lineages with broad admixture across geographic hybrid zones, which could potentially represent distinct species. However, while limited dispersal promotes substantial isolation by distance, extreme microhabitat specificity appears to yield stabilizing selection on ecologically mediated phenotypes. As a result, climatic cycles promote recurrent contact between lineages that are not adaptively differentiated and therefore experience repeated bouts of high migration and introgression through time. This leads to a unified, single species with deeply divergent phylogeographic lineages that nonetheless do not appear to represent incipient species.

## Discerning structure versus speciation in phylogeographic analysis of Seepage Salamanders (*Desmognathus aeneus*) using demography, environment, geography, and phenotype

R. Alexander Pyron<sup>1,2\*</sup>, Anvith Kakker<sup>3</sup>, David A. Beamer<sup>4</sup>, Kyle A. O'Connell<sup>2,5</sup>

<sup>1</sup>*Department of Biological Sciences, The George Washington University, 2023 G St. NW, Washington, DC 20052; rpyron@gwu.edu; <https://orcid.org/0000-0003-2524-1794>*

<sup>2</sup>*Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560*

<sup>3</sup>*Thomas Jefferson High School for Science and Technology, 6560 Braddock Rd, Alexandria, VA 22312*

<sup>4</sup>*Office of Research, Economic Development and Engagement, East Carolina University, 209 East Fifth Street, Greenville, NC 27858; <https://orcid.org/0000-0003-0796-274X>*

\*Corresponding author

**Abstract:** Numerous mechanisms drive ecological speciation, including isolation by adaptation, barrier, distance, environment, hierarchy, and resistance. These promote genetic and phenotypic differentiation of local populations, formation of phylogeographic lineages, and ultimately, completed speciation via reinforcement. In contrast, it is possible that similar mechanisms might lead to lineage cohesion through stabilizing rather than diversifying ecomorphological selection and the long-term persistence of population structure within species. Processes that drive the formation and maintenance of geographic genetic diversity while facilitating high rates of migration and limiting phenotypic divergence may thereby result in population structure that is not accompanied by divergence towards reproductive isolation. We suggest that this framework can be applied more broadly to address the classic dilemma of “structure versus speciation” when evaluating phylogeographic diversity, unifying population genetics, species delimitation, and the underlying study of speciation. We demonstrate one such instance in the Seepage Salamander (*Desmognathus aeneus*) from the southeastern United States. Recent studies estimated up to 6.3% mitochondrial divergence and 4 phylogenomic lineages with broad admixture across geographic hybrid zones, which could potentially represent distinct species. However, while limited dispersal promotes substantial isolation by distance, extreme microhabitat specificity appears to yield stabilizing selection on ecologically mediated phenotypes. As a result, climatic cycles promote recurrent contact between lineages that are not adaptively differentiated and therefore experience repeated bouts of high migration and introgression through time. This leads to a unified, single species with deeply divergent phylogeographic lineages that nonetheless do not appear to represent incipient species.

**Keywords:** Phylogeographic structure, ecological speciation, ecomorphological adaptation, gene flow, migration rates, Pleistocene climatic fluctuations

## Introduction

The presence of substantial population structure within many if not most taxa is well-documented empirically (Slatkin 1987; Palumbi et al. 1997), and delimiting species in such instances has become a focal topic in systematics (Hey and Pinho 2012; Sukumaran and Knowles 2017). A variety of processes can generate geographic genetic structure such as local environmental adaptation and isolation by distance (Duminil et al. 2007; Leffler et al. 2012), accumulating deep divergence between populations over time and resulting in multiple phylogeographic lineages within cohesive species (Rissler et al. 2006; Soltis et al. 2006). If these divergences persist and are not resorbed (Rosenblum et al. 2012), the eventual outcome is generally speciation (Dynesius and Jansson 2014), wherein populations diverge sufficiently to acquire independent evolutionary trajectories characterized by reproductive isolation and low rates of hybridization (Singhal and Moritz 2013). Completed speciation is often characterized by adaptive ecological, genetic, and phenotypic differentiation that reinforce species boundaries via reproductive isolation and selection against hybrids over time (Coyne and Orr 2004).

In contrast, substantial population structure can originate within species that is nonetheless held in check by high rates of gene flow (Bohonak 1999) and is consequently not indicative of incipient, incomplete, or ongoing speciation (see Huang and Knowles 2016), but is instead a long-term, stable, or persistent endpoint on its own (Yang et al. 2019). Furthermore, the degree and rate of population divergence may or may not be correlated to rate of completed speciation, depending on geography and connectivity (Kisel et al. 2010; Harvey et al. 2017; Singhal et al. 2018b, 2022; Burbrink et al. 2023). Secondary contact can even result in lineage fusion (Maier et al. 2019) and speciation reversal (Kearns et al. 2022), with subsequent gain of structure (Frei et al. 2022). Consequently, population structure may represent the outcome of various processes yielding geographic genetic diversity within species that nonetheless do not reflect speciation trajectories (Cutter and Gray 2016; Sukumaran and Knowles 2017).

Genetic structure in species may therefore consist of divergence that is neutral (e.g., isolation by barrier, distance, or hierarchy; Baptestini et al. 2013) or adaptive (e.g., ecological selection resulting in isolation

by adaptation or environment; Nosil et al. 2008; Wang and Bradburd 2014). Such isolation (Zbinden et al. 2022; Moreno-Contreras et al. 2023) can consequently produce regional divergence that is bound together by migration over time and space (Garnier et al. 2004; Sexton et al. 2014). This may represent incomplete speciation (Nosil et al. 2009) or ephemeral differentiation that is ultimately resorbed (Rosenblum et al. 2012). Protracted scenarios of speciation-like outcomes can also occur wherein evolutionarily distinct lineages nevertheless exhibit gene flow in primary or secondary contact (Smadja and Butlin 2011; Pardo-Diaz et al. 2012; Burbrink et al. 2021). Complex landscapes may therefore produce substantial local differentiation (Brauer et al. 2018; Nali et al. 2020), interacting with climatic cycles to produce recurring episodes of divergence (Fitzpatrick et al. 2009) and secondary contact with ongoing gene flow (Schield et al. 2019; O’Connell et al. 2021).

Consequently, deep intraspecific divergence of genetic lineages can be generated by a variety of geographic and ecological processes that are either i) part of a generalized trajectory of divergence towards speciation, or ii) population structure that is nonetheless unified by gene flow (Avice 2000; Coyne and Orr 2004). Therefore, demographic model selection and tests of isolation-by-distance (Jackson et al. 2017) can be instrumental for differentiating structure versus speciation (Sukumaran and Knowles 2017). When the former is produced by landscape-scale processes that are held in check by ongoing migration (Seeholzer and Brumfield 2018), this is detectable by a variety of methods (Carstens et al. 2022). In contrast, speciation is demonstrated by reduced migration (below thresholds of isolation by distance or environment) between ecologically, geographically, and phenotypically divergent lineages (Burbrink et al. 2021).

Finally, integrated analyses of phenotype that test for the correlated ecomorphological signature of phylogeographic lineage divergence are also crucial for testing speciation (Zamudio et al. 2016), even in putatively cryptic taxa (Singhal et al. 2018a). Consequently, multi-modal analyses are needed to discern an accurate picture in many cases (Bertl et al. 2018); the data and methods needed to test structure versus speciation will often be the same as those needed to identify the mode, mechanism, and process of ecological, geographic, genetic, and phenotypic divergence (Pyron et al. 2023). The fundamental test is therefore whether the processes driving the accumulation of geographic genetic structure (e.g., IBD/IBE) are implicated in diversifying ecomorphological selection promoting divergence between the incipient phylogeographic lineages, or whether other mechanisms simultaneously foster high levels of migration between these populations sufficient to promote lineage cohesion and maintain species identity (Kisel et al. 2010; Seeholzer and Brumfield 2018; Freedman et al. 2023; Prates et al. 2023).

Here, we evaluate “structure” versus “speciation” in the Seepage Salamander (*Desmognathus aeneus*) using tests of isolation-by-distance, automated demographic-model selection, species distribution modeling, genotype-environment and -phenotype association, and morphometric divergence. Seepage Salamanders are miniaturized (~3–5cm total length), terrestrial, direct-developing lungless salamanders (Plethodontidae) from the southeastern United States, primarily from the Blue Ridge of Georgia and North Carolina (Harrison 1992). Given their size, their home range and dispersal abilities are miniscule, and they are therefore highly susceptible to landscape-level genetic fragmentation over ecological timescales. They also exhibit strong ecological specificity for seepage habitats associated with the headwaters of high-gradient creeks and streams. Interestingly, they are also found in analogues of this microhabitat at low elevations in dispersed, isolated pockets across the Piedmont, Southeastern Plains, Ridge and Valley, and Southwestern Appalachians in central Alabama and western Georgia, some of which were not discovered until the 2010’s (Graham et al. 2012).

Recent range- and genome-wide results indicate up to 6.3% mitochondrial divergence between four deeply divergent but spatially adjacent phylogenomic lineages with extensive admixture (Beamer and Lamb 2020; Pyron et al. 2020, 2022). Given their broad but fragmented geographic range, we suggest that Seepage Salamanders occupied a continuous distribution across the southeastern U.S. during cooler glacial periods. Similarly, we hypothesize that their microhabitat specificity results in stabilizing ecological selection and species cohesion despite landscape genetic variation. Finally, we expect morphometric similarity to reflect this ecological uniformity. We find that the evolutionary history of the group is characterized by deep lineage

divergence that is counterbalanced by episodes of increased migration and a lack of phenotypic differentiation; “structure” rather than “speciation.” It is difficult to pinpoint the exact nature of geographic genetic diversity (e.g., IBD/IBE) given the complex ecological nature of their microhabitat adaptations, apparently involving hydrological, edaphic, floristic, and climatic interactions. Nevertheless, the frequency of estimated migration and observed admixture in individuals suggests that reinforcement is not operating to select against hybridization between the genetically differentiated populations. Formal tests of the dichotomy between structure and speciation should be employed when this distinction is consequential in phylogeography.

## Materials and Methods

### *Genetic Data*

We sampled 56 individuals from 31 localities (Appendix S1; Fig. 1) from the known range of the species, including the recently discovered populations in the Piedmont and Coastal Plain of Georgia (Graham et al. 2012). This includes isolated localities in central and western Alabama, including near the type locality of *Desmognathus chermocki* Bishop and Valentine, 1950, a junior synonym of *D. aeneus* Brown and Bishop, 1947 after Chermock (1952). We generated SNP data using genotype-by-sequencing (GBS; Elshire et al. 2011) with the ApeKI enzyme at the University of Wisconsin-Madison Biotechnology Center. We sequenced libraries on a NovaSeq6000 (Illumina Inc.), with coverage ranging from 2.3–14.6 (mean 10.2) million reads per sample. We optimized assemblies in ipyrad v.0.9.87 (Eaton and Overcast 2020) using the criteria proposed by Ilut et al. (2014) and McCartney-Melstad et al. (2019). These yielded an optimal threshold of 97% similarity for clustering on an initial sequencing run of 33 individuals (Fig. S1), which we retained for the full analysis of 56. We enforced a threshold of >60% presence per locus, equal to a minimum of 34/56 individual coverage across samples. The raw assembly yielded 13,091 SNPs with 26,593 alleles from 2,132 loci. After filtering for missing data (>33%) and singleton alleles (Linck and Battey 2019), we retained 2,281 SNPs with 4,661 alleles. A group of seven samples had <50% coverage and were therefore dropped for the conStruct, IBD, and RDA analyses (see below), which are sensitive to missing data. Part of these analyses were carried out on the GW HPC *Pegasus* cluster (MacLachlan et al. 2020).

### *Genetic Clustering*

We estimated genetic clusters using discriminant analysis of principal components (DAPC; Jombart et al. 2010) and sparse non-negative matrix factorization (sNMF; Frichot et al. 2014) on the SNP matrix. For DAPC in the R package ‘adegenet’ (Jombart 2008), we chose the optimal number of clusters by minimum Bayesian Information Criterion (BIC) and optimal assignment using DAPC cross-validation (Fig. S2–4). For sNMF in ‘LEA’ (Frichot and François 2015), we chose the regularization factor  $\alpha$  and the number of clusters  $K$  using minimum median cross-entropy from 100 replicates (Fig. S5–6). For a preliminary test of IBD, we performed a spatially-aware analysis of population clustering and admixture using the ‘conStruct’ package (Bradburd et al. 2018), which accounts for geographic genetic variation using spatial layers to represent clusters that account for allelic covariance attributed to continuous (e.g., IBD) versus discrete (e.g., barriers) sources (Bradburd and Ralph 2019). We tested 1–4 layers with 4 chains of 250k generations. We took a layer to be significant when it accounted for more than 20% of the allelic covariance matrix (Fig. S7).

### *Linear Morphometrics*

We modified our recent approaches (Pyron and Beamer 2023) based on Bingham et al. (2018)’s strategy for multivariate assessment of salamanders. We recorded 17 measurements at 0.01mm precision for 89 specimens from the 31 sites: SVL (snout-vent length), TL (tail length), AG (axilla-groin length), CW (chest width), FL (femur length), HL (humerus length), SG (snout-gular length), TW (tail width at rear of vent), TO (length of third toe), FI (length of third finger), HW (head width), ED (eye diameter), IN (internarial distance), ES (eye-snout distance), ON (orbito-narial distance), IO (inter-orbital distance), and IC (inter-canthal distance). We implemented the general recommendations and statistical approaches of (Onn and Grismer 2021; Chan and Grismer 2022) for linear morphometrics, allometric size-correction, and quantitative analyses. We employed the “common within-group pooling” strategy, wherein distinct populations or localities within a species (e.g., geographic lineages) have separate slopes but a shared intercept representing

overall body-size for the whole species. We tested both pooled-group and multi-species approaches; they were qualitatively identical, and we therefore present the more conservative pooled-group analyses. After allometric size-correction, we performed Linear-Discriminant Analysis (LDA) to examine phenotypic differentiation of populations. We identified variables with the heaviest loading across the discriminant axes as a preliminary indicator of the morphological trait with the greatest importance in differentiating populations.

### *Species Distribution Modeling*

As a preliminary estimate of habitat suitability given present-day climatic conditions, we estimated species distribution models using the ‘maxent’ algorithm in the ‘dismo’ package (Hijmans et al. 2017). We included the 31 GBS localities from our dataset and the 59 AHE localities from Pyron et al. (2022c). As an initial set of predictors, we included 37 variables: WorldClim (Fick and Hijmans 2017), ENVIREM (Title and Bemmels 2018), and Level IV ecoregions (Omernik and Griffith 2014). These were sampled at 30s resolution and re-projected to North America Albers Equal Area Conic (ESRI:102008). We used the ‘corSelect’ function in the ‘fuzzySim’ package (Barbosa 2015) to remove multicollinear predictors based on variance inflation factor (VIF). The retained set 6 BIOCLIM and 3 ENVIREM variables and the Level IV Ecoregions; the full list is given in the SI. We optimized a maxent model using the 90 localities and 1,000 background pseudo-absence points from a 250km radius. We then projected this model as a three-level binary prediction under the equate entropy of thresholded and original distributions, maximum training sensitivity plus specificity, and equal training sensitivity and specificity to evaluate varying degrees of potential occupancy across the range (Liu et al. 2015).

### *Isolation By Distance*

As an initial test of IBD, we implemented various classical statistics based on  $F_{st}$  values (Slatkin 1993; Rousset 1997; Hutchison and Templeton 1999; Diniz-Filho et al. 2013). We calculated pairwise  $F_{st}$  using the approach of Weir and Cockerham (1984) for individuals, populations, and lineages. For individuals, we used a Mantel test with 999 permutations to evaluate significance of the overall relationship between geographic and genetic distance. For populations, we examined between- and within-lineage slopes to estimate the relative prominence of local versus regional structure. Finally, we calculated overall fixation between lineages as a basic estimator for potentially species-level divergence (Roux et al. 2016). For individual and population-based analyses, we used linearized  $F_{st} = F_{st} / (1 - F_{st})$ .

As an additional test of both IBD and isolation by environment (IBE), we used generalized dissimilarity modeling (GDM) in the R package ‘gdm’ (Ferrier et al. 2007; Mokany et al. 2022). We used the pairwise  $F_{st}$  matrix by site as our measure of genetic dissimilarity, and the 9 continuous ecological variables, elevation, and geographic distance (lat/long; see SI for list) as the environmental predictors. We fit a GDM to these data and assessed variable inclusion using estimated coefficients greater than zero, and variable importance and significance using 1,000 permutations. We also evaluated model significance and percent deviance explained.

### *Demographic Modeling*

We optimized a three-population demographic model based on the site-frequency spectrum (SFS) using the genetic algorithm ‘GADMA’ (Noskova et al. 2020, 2023) based on the ‘moments’ engine (Jouanous et al. 2017). We first down projected the VCF file from ipyrad to a three-dimensional SFS of [27,9,21] from (28,8,20) diploid individuals from the NTA, COH, and CHR lineages (see Results) using ‘easySFS’ (<https://github.com/isaacovercast/easySFS>; Gutenkunst et al. 2009). Down projection was based on the allelic sample size that maximized the number of segregating sites per lineage. Mutation rate (per site, per generation) is not generally known in amphibians, so we used the mean rate for fishes of 5.97e-09; these values are relatively similar across vertebrates (e.g., mean for mammals of 7.97e-09; Bergeron et al. 2023). The generation time for *Desmognathus aeneus* is ~3 years (Bruce 2022). Specifying mutation rate and sequence length allow inference of  $\vartheta$ , and generation time allows re-scaling to thousands of years. We allowed asymmetric migration, unlinked SNPs with no outgroup, and an initial and final structure with one time-period per divergence (i.e., 1,1,1). We used 120 repeated optimizations in GADMA to arrive at the final models (Fig. S8).

## Adaptive Potential

Finally, we performed a preliminary screen for the potential signature of multi-locus adaptation and IBE using redundancy analysis (RDA; Forester et al. 2018; Capblancq and Forester 2021). We evaluated Genotype-Environment Associations (GEA) and Genotype-Phenotype Association (GPA) to identify sets of candidate loci that might be adaptively differentiated in response to both sets of predictors. We then estimated whether these sets of loci overlap to determine any signal of concerted selection or adaptation linking ecomorphological divergence (Zhang et al. 2019; Carvalho et al. 2021; Valette et al. 2023). While the lack of an annotated genome and identification of the GBS loci limits subsequent inference, the genome-wide nature of the markers should capture at least some variation related to adaptive ecological and phenotypic differentiation, if present. We used the 9 climatic variables from the GDM analysis for the 49 samples with low missing data, and 14 phenotypic measurements for 40 of those samples that matched between the two datasets. We dropped IN, TL, and ES from the trait data, as they had  $\sim 0$  loadings in the LDA between populations (Fig. 2).

## Results

### Distribution

The estimated distribution from the Maxent model based on 90 presence localities encompassing nearly all known, extant populations essentially matches the range of the species, generally with high predicted suitability (Fig. 1). Small regions of high suitability in northwest Georgia and the Cumberland Plateau of south-central Tennessee and northeastern Alabama may represent favorable extralimital habitat but should be intensively surveyed for the possibility of historically distributed populations that may have escaped previous notice. Of the 10 variables included in the model, only PET of the Wettest Quarter had a substantial contribution to regularized gain (63%); followed by PET of the Warmest Quarter (10%); and Level IV Ecoregions, BIO4 – Mean Diurnal Range, and BIO18 – Precipitation of the Warmest Quarter (7%). All remaining variables contributed less than 5%, suggesting low overall importance.

### Structure, Demographics, and Phenotype

Results from DAPC and sNMF are highly concordant in estimating three major lineages (Fig. 3). These correspond primarily to populations in the Piedmont (‘CHR’), Cohutta Mountains (‘COH’) in north-central Georgia, and Nantahala mountains in southwestern North Carolina and adjacent north-east Georgia (‘NTA’). In contrast, conStruct estimates two distinct lineages (CHR and NTA+COH) when accounting for spatial distance, with the second layer accounting for  $\sim 80\%$  of allelic covariance (Fig. S7). Estimates accounting for admixture in sNMF and conStruct suggest widespread introgression at the contact zones between these lineages (montane/Piedmont in the latter), with substantial genomic ancestry ( $>10\%$ ) from each lineage inferred for multiple individuals from multiple populations across most of the species’ range.

The GADMA estimates converged on a model with relatively small ancestral population sizes ( $\sim 240k$ ) originating in the mid-Pleistocene ( $\sim 1.2Ma$ ), with recent exponential increases ( $\sim 3\text{--}19\times$ ) during the latest glacial cycles  $\sim 630Ka$ –present (Fig. 3). Migration was very high between the two ancestral lineages (estimated at the upper bound of  $2Nm = 20$ ) and decreased towards the present; remaining  $>1$  for NTA $\leftrightarrow$ COH (1.24 and 1.85), CHR $\rightarrow$ NTA (1.40), and CHR $\rightarrow$ COH (1.70). Populations are not phenotypically differentiated based on the 17 linear morphometric measurements taken here (Fig. 2). Specimens from all 3 lineages overlap in linear-discriminant space, where the heaviest-loading size-corrected variable of trunk length (AG; axilla-groin) does not differ significantly between them. Overall, these results are congruent with the hypothesis of a single deeply structured species, wherein the constituent phylogeographic lineages are not phenotypically differentiated and maintain migration sufficient to overcome drift, local adaptation, and divergent ecological selection.

### Distance and Environment

A basic Mantel test of geographic and genetic distance was strongly significant ( $r = 0.56$ ,  $P < 0.001$ ) based on 999 permutations. Pooled across sites, pairwise  $F_{st}$  for NTA, COH, and CHR equal 0.16, 0.18, and 0.19, which even under simplistic models (Wright 1931; Slatkin 1993) suggests 1.04–1.35 migrants per

generation between the three major lineages (see Mills and Allendorf 1996). Examination of linearized  $F_{st}$  against geographic distance reveals an apparently strong linear relationship  $<50\text{km}$ , which levels off to  $\sim 0$  after  $100\text{km}$  (Fig. 4). This is driven by strong within-lineage IBD for the geographically restricted, montane Cohutta and Nantahala lineages and little to no between-lineage relationship or within the Piedmont lineage. These results were corroborated by the GDM, which dropped elevation, BIO9, and BIO13 based on a  $\sim 0$  sum of coefficients, and only estimated geographic distance as significant ( $P = 0.04$ ) with a similar exponential plateau after  $\sim 50\text{--}100\text{km}$  (Fig. 3–4). These findings mirror the overall low impact of most climatic variables in the SDM and suggest a relatively minimal role for IBE in geographic genetic diversification for any measured predictors. The overall model was significant at  $P < 0.0001$  and explained 52% deviance, an ordinary amount within the expected range of several past studies (see Mokany et al. 2022).

### *Multi-Locus Adaptation*

The RDA using 7 environmental variables identified 77 SNPs significantly correlated with 4 of them (Fig. 5): months  $> 10$  degrees C (37 loci), BIO4 – Temperature Seasonality (26), PET of the warmest quarter (12), and BIO18 – Precipitation of the Warmest Quarter (2). The phenotypic analysis estimated 30 significant associations with 9 variables: ON and FL (6 each); TW and TO (4); and SVL, IC, HL, ED, and CW (2). These primarily reflect body size, head shape, and limb length. Of the significant loci, 18 are shared between the environmental and phenotypic predictors, with 12 related to months  $> 10$  degrees C and limb length or head shape. Of those remaining, 2 SNPs are related to PET of the Wettest Quarter and tail width, and 4 are related to Temperature Seasonality and head shape. Consequently, we conclude that there is a limited but non-zero degree of adaptive genetic differentiation with respect to ecomorphology along axes of temperature, precipitation, and phenotypic robustness. Given the lack of a suitable reference genome, we cannot pinpoint the physical location of these sites.

## **Discussion**

### *Phylogeographic Divergence*

Our recent phylogenomic dataset sampled 59 individuals and populations and estimated four geographic lineages (Pyron et al. 2022), but with fewer samples outside of the Blue Ridge mountains and including only 233 loci. Previous mitochondrial datasets indicated substantial population divergence in the species (Beamer and Lamb 2020; Pyron et al. 2020), with uncorrected ‘ $p$ ’ distances in the COI barcode locus of 1.4–6.3%. Here, we recover three lineages rather than four, corresponding to two major mountain ranges (Cohutta and Nantahala) and the expansive Piedmont and associated populations. Accounting for IBD when estimating population structure in conStruct suggests only two major clusters (Piedmont versus montane). Overall, lineages of Seepage Salamander are deeply divergent ( $\sim 1.2\text{Ma}$ ,  $\sim 4.7\%$  mitochondrially), but exhibit extensive admixture both at their contact zones and deep into the core geographic range of each set of populations. The between-lineage  $F_{st}$  values border on “meaningful” significance (usually considered  $\sim 0.15$ ) for population differentiation (Hartl and Clark 2007), but generally fall outside or on the low end of the “grey zone” of genomic divergence ( $F_{st} > 0.2, D_a > 0.01$ ) indicative of speciation across a wide variety of animals (Roux et al. 2016). Recent migration rates inferred from  $F_{st}$  and estimated by GADMA are  $> 1$  migrant/generation between all three lineages, suggesting relatively high and constant rates of gene flow through time.

The overall relationship between geographic and genetic distance between individuals, populations, and lineages matches the classic expectation of a “Case IV” scenario resulting from a lack of regional equilibrium (Hutchison and Templeton 1999). This pattern is typically driven by differing scale-dependent effects of gene flow and drift influencing population structure (Twyford et al. 2020), where migration has greater impacts over shorter distances and drift predominates past a threshold determined in part by habitat connectivity and migration range (van Strien et al. 2015). The lack of between-lineage IBD likely reflects the long-term impact of mountains and refugia in structuring local populations, explaining the presence of very close yet very different populations in the Blue Ridge. This also rejects IBA, as the strength of between-population IBD is less than within-population, indicating no acceleration of genetic divergence due to local adaptation in the Blue Ridge and Piedmont (Freedman et al. 2023). Correspondingly, the relatively recent expansion

of the Piedmont lineages during the late Pleistocene explains the weakness of IBD within the most widely distributed lineage, as “Case I” dynamics (i.e., “pure” IBD) require relatively long timescales to become apparent. This again underscores the dynamic nature of these processes, and the attendant capacity of related patterns to shift over time.

The RDA results confirm the hypothesis that there is some apparent degree of ecological adaptation related to temperature and precipitation differences between montane and Piedmont populations (Miranda et al. 2023), but they are not abundant in the genome. Furthermore, there is a non-zero but limited degree of correlated adaptive divergence along ecological gradients and related phenotypic axes. Robustness in terms of size, limb length, and head shape (length and width) is likely related to both desiccation tolerance in warmer environments (Baken et al. 2020) and differing relative pressures of terrestriality (Ledbetter and Bonett 2019). That these variables show some degree of correlated adaptive differentiation between montane and Piedmont environments is not surprising, but it does not result in overall phenotypic divergence between lineages or apparently act to limit hybridization and gene flow. We consider our limited number of predictors in both analyses to offer a tradeoff between potential false positives and false negatives (Forester et al. 2018). Our sampling is not detailed enough to perform cluster-specific analyses (Carvalho et al. 2021), but a better-annotated molecular dataset could facilitate this in the future to detect microgeographic variation and adaptation.

The SDMs along with other natural-history observations (Harrison 1992; Graham et al. 2012) reinforce the strong influence of Level IV Ecoregions on the distribution of this species, particularly outside of the Blue Ridge Mountains. Similar results were seen in the approximately co-distributed congener *Desmognathus cheaha* (Pyron et al. 2023). We interpret this to indicate floristic or other associations (edaphic, hydrological, etc.) with seepage environments not captured by climatic variation and liable to change dramatically over ecological timescales (A. Lee-Yaw et al. 2022). Our fieldwork underscores the extreme microhabitat specificity involved in locating this species, but we have not yet pinpointed the exact determinants of what constitutes suitable habitat. For this reason, it is difficult to test hypotheses about historical distributions with paleoclimatic modeling given the lack of a mechanistic model and historical data layers that incorporate the relevant factors, which are still mostly unknown. Similarly, we did not consider modern estimated effective migration surfaces (Petkova et al. 2016) to be particularly relevant for investigating these historical processes, which shift rapidly through space and time.

Consequently, we hypothesize that structure in this species arises primarily from stabilizing rather than diversifying ecomorphological selection. This results from a high degree of ecological specificity to a highly specific microhabitat consisting of headwater seepages, moist leaf litter, and various moss species. These salamanders are almost never observed outside of this exact ecosystem, typically within the range of a few meters around spring heads or ravine streams. Yet, increases in suitable habitat during periods of cooling and expansion out of refugia appear to drive rapid geographic occupancy across a large expanse of the Piedmont and associated ecoregions, which is quickly fragmented during interglacial periods. This adaptive stasis and limited dispersal both drives IBD in geographically proximate areas but contributes to lineage cohesion across climatic cycles when these locally adapted lineages are brought into contact out of climatically proximate refugia during glacial cycles. Similar processes appear to be operating in other salamander systems in Mexico (Velo-Antón et al. 2013), suggesting analogous montane processes linking tropical and temperate dynamics in landscape genetics and lineage formation. This is essentially an extension of Janzen’s hypothesis (Janzen 1967; Muñoz and Bodensteiner 2019; Wishingrad and Thomson 2023), wherein behavior, ecology, and phenotype interact to drive local adaptation, constrain ecomorphological divergence, and promote lineage cohesion. Interestingly in this case, these processes ultimately appear to foster phylogeographic diversification while constraining speciation trajectories.

### *Structure Versus Speciation*

Speciation and structure may occasionally present superficially similar patterns but can theoretically be distinguished by mechanistic hypothesis testing of the underlying processes. Furthermore, the empirical “gray zone” of speciation is oftentimes fairly binary and sharply demarcated (Roux et al. 2016; Dufresnes et

al. 2021). When relevant axes such as genomic divergence, ecological adaptation, migration, and phenotypic differentiation are accounted for, an integrative evaluation of speciation hypotheses can discern between the two (Freedman et al. 2023; Prates et al. 2023; Pyron et al. 2023). When divergent ecomorphological selection is driving ecological speciation, we would expect correlated genomic signatures of genetic and phenotypic divergence along environmental axes and reduced rates of migration (e.g., Orsini et al. 2013). In contrast, limited phenotypic diversity and high rates of migration can allow for the long-term persistence of deeply divergent phylogeographic lineages and geographic genetic diversity within genomically cohesive single species over long periods of time (e.g., Postma and van Noordwijk 2005; Kremer et al. 2012; Fitzpatrick et al. 2015).

Correspondingly, we emphasize that the type of deep population structure we describe here is not necessarily simply a point along the speciation continuum (Stankowski and Ravinet 2021) but a distinct pattern resulting from similar processes. The forces we hypothesize to be driving genetic divergence (microhabitat specificity) also limit ecomorphological diversification and promote recurrent secondary contact during glacial cycles as populations expand and contract in and out of physical proximity. We suggest that such mechanisms can generate substantial amounts of stable or enduring phylogeographic diversity that is not intrinsically on a path to speciation (Huang 2020), rather than simply an intermediate or incipient phase along the speciation continuum (Dufresnes et al. 2019). Consequently, this suggests a model in which the underlying processes that generate geographic genetic diversity are similar, but the outcomes of stable or persistent population structure versus ecological speciation differ (Nosil et al. 2009). In the latter, a gray zone of incipient or incomplete speciation along the continuum of ecomorphological divergence may also sometimes be observed (Huang and Knowles 2016).

Ecological speciation may occur via many routes, but any mechanism that promotes high rates of continued migration may act against those processes and maintain population structure without speciation. This is important, given that even extremely low rates of migration (i.e.,  $\ll 1$ /generation) can lead to the substantial signal of introgression across the genome (Jiao et al. 2020; Jiao and Yang 2021). While it may not necessarily be easy to distinguish complex spatial and temporal patterns of structure versus speciation (DeRaad et al. 2022; Prates et al. 2023), we provide a framework here for attempting to do so in future studies. Combining analyses of migration, spatial demographics, and adaptive ecomorphological divergence allows us to test for structure versus speciation, conduct species delimitation, and identify potential underlying mechanism and process of speciation itself (Maier et al. 2019; Smith and Carstens 2020; Burbrink et al. 2021; O’Connell et al. 2021; Freedman et al. 2023; Pyron et al. 2023).

## Conclusions

Mechanisms including isolation by adaptation, barrier, environment, hierarchy, and resistance can lead to local genetic diversity and the accumulation of deeply divergent phylogeographic lineages within species. If divergent ecological selection acts to enhance population divergence along any number of axes such as phenotype, microhabitat, or behavior to reduce migration and increase reproductive isolation, then speciation is the natural and expected outcome. In contrast, multiple processes may counter this, enhancing rates of migration and gene flow and limiting ecomorphological diversification despite the accumulation of geographic genetic lineages. This may lead to long-term prevalence of population structure within species that is not necessarily tied to incipient, incomplete, or ongoing speciation, but is instead a stable or persistent endpoint held in check by migration. We demonstrate one such example here in Seepage Salamanders, *Desmognathus aeneus*. Geographic lineages dating back to the mid-Pleistocene are nonetheless unified by high levels of migration through time, while stabilizing ecomorphological selection arising from extreme microhabitat specificity apparently limits potential ecological speciation. This conclusion is reinforced by the limited but non-zero genomic signature of correlated genetic and phenotypic divergence along axes of temperature and precipitation between montane and lowland regions. The framework presented here clarifies many of the theoretical expectations and offers potential empirical tests for the previously ambiguous definition of structure versus speciation in the quantification of phylogeographic lineage diversity.

## Acknowledgments

This research was funded in part by the following awards: US NSF grant DEB-1655737 and GW UFF to RAP, DEB-1656111 to DAB, and an NMNH Peer Award to KAO. This work was completed in part with resources provided by the High Performance Computing Cluster at The George Washington University (Information Technology, Research Technology Services), and the Smithsonian Institution High Performance Cluster (<https://doi.org/10.25572/SIHPC>). We thank S. Graham and C. Kelehear for field assistance.

## References

- Avice J.C. 2000. *Phylogeography: the history and formation of species*. Cambridge, Mass: Harvard University Press.
- Baken E.K., Mellenthin L.E., Adams D.C. 2020. Macroevolution of desiccation-related morphology in plethodontid salamanders as inferred from a novel surface area to volume ratio estimation approach. *Evolution*. 74:476–486.
- Baptistini E.M., de Aguiar M.A.M., Bar-Yam Y. 2013. Conditions for neutral speciation via isolation by distance. *J Theor Biol*. 335:51–56.
- Barbosa A.M. 2015. fuzzySim: applying fuzzy logic to binary similarity indices in ecology. *Methods in Ecology and Evolution*. 6:853–858.
- Beamer D.A., Lamb T. 2020. Towards rectifying limitations on species delineation in dusky salamanders (*Desmognathus*: Plethodontidae): An ecoregion-drainage sampling grid reveals additional cryptic clades. *Zootaxa*. 4734:1–61.
- Bergeron L.A., Besenbacher S., Zheng J., Li P., Bertelsen M.F., Quintard B., Hoffman J.I., Li Z., St. Leger J., Shao C., Stiller J., Gilbert M.T.P., Schierup M.H., Zhang G. 2023. Evolution of the germline mutation rate across vertebrates. *Nature*.:1–7.
- Bertl J., Ringbauer H., Blum M.G.B. 2018. Can secondary contact following range expansion be distinguished from barriers to gene flow? *PeerJ*. 6:e5325.
- Bingham R.E., Papenfuss T.J., Lindstrand L., Wake D.B. 2018. Phylogeography and Species Boundaries In the *Hydromantes shastae* Complex, With Description of Two New Species (Amphibia; Caudata; Plethodontidae). *Bulletin of the Museum of Comparative Zoology*. 161:403–427.
- Bishop S.C., Valentine B.O. 1950. A New Species of *Desmognathus* from Alabama. *Copeia*. 1950:39.
- Bohonak A.J. 1999. Dispersal, Gene Flow, and Population Structure. *The Quarterly Review of Biology*. 74:21–45.
- Bradburd G.S., Coop G.M., Ralph P.L. 2018. Inferring Continuous and Discrete Population Genetic Structure Across Space. *Genetics*. 210:33–52.
- Bradburd G.S., Ralph P.L. 2019. Spatial Population Genetics: It’s About Time. *Annual Review of Ecology, Evolution, and Systematics*. 50:427–449.
- 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*. 27:3484–3497.
- Brown W.C., Bishop S.C. 1947. A New Species of *Desmognathus* from North Carolina. *Copeia*. 1947:163.
- Bruce R.C. 2022. Size and Cycle in Dusky Salamanders. *Journal of Herpetology*. 56:444–453.
- Burbrink F.T., Gehara M., McKelvy A.D., Myers E.A. 2021. Resolving spatial complexities of hybridization in the context of the gray zone of speciation in North American ratsnakes ( *Pantherophis obsoletus* complex). *Evolution*. 75:260–277.

- Burbrink F.T., Ruane S., Rabibisoa N., Raselimanana A.P., Raxworthy C.J., Kuhn A. 2023. Speciation rates are unrelated to the formation of population structure in Malagasy gemsnakes. *Ecology and Evolution*. 13:e10344.
- Capblancq T., Forester B.R. 2021. Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods in Ecology and Evolution*. 12:2298–2309.
- Carstens B.C., Smith M.L., Duckett D.J., Fonseca E.M., Thomé M.T.C. 2022. Assessing model adequacy leads to more robust phylogeographic inference. *Trends in Ecology & Evolution*.
- Carvalho C.S., Forester B.R., Mitre S.K., Alves R., Imperatriz-Fonseca V.L., Ramos S.J., Resende-Moreira L.C., Siqueira J.O., Trevelin L.C., Caldeira C.F., Gastauer M., Jaffé R. 2021. Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration. *Molecular Ecology Resources*. 21:44–58.
- Chan K.O., Grismer L.L. 2022. GroupStruct: An R Package for Allometric Size Correction. *Zootaxa*. 5124:471–482.
- Chermock R.L. 1952. A key to the amphibians and reptiles of Alabama. University, Ala.
- Coyne J.A., Orr H.A. 2004. *Speciation*. Oxford, New York: Oxford University Press.
- Cutter A.D., Gray J.C. 2016. Ephemeral ecological speciation and the latitudinal biodiversity gradient. *Evolution*. 70:2171–2185.
- DeRaad D.A., Applewhite E.E., Tsai W.L.E., Terrill R.S., Kingston S.E., Braun M.J., McCormack J.E. 2022. Hybrid zone or hybrid lineage: a genomic reevaluation of Sibley’s classic species conundrum in *Pipilo towhees*. *Evolution*.:qpac068.
- Diniz-Filho J.A.F., Soares T.N., Lima J.S., Dobrovolski R., Landeiro V.L., de Campos Telles M.P., Rangel T.F., Bini L.M. 2013. Mantel test in population genetics. *Genet Mol Biol*. 36:475–485.
- Dufresnes C., Brelford A., Jeffries D.L., Mazepa G., Suchan T., Canestrelli D., Niecieza A., Fumagalli L., Dubey S., Martínez-Solano I., Litvinchuk S.N., Vences M., Perrin N., Crochet P.-A. 2021. Mass of genes rather than master genes underlie the genomic architecture of amphibian speciation. *PNAS*. 118.
- Dufresnes C., Strachinis I., Suriadna N., Mykytynets G., Cogălniceanu D., Székely P., Vukov T., Arntzen J.W., Wielstra B., Lymberakis P., Geffen E., Gafny S., Kumlutaş Y., Ilgaz Ç., Candan K., Mizsei E., Szabolcs M., Kolenda K., Smirnov N., Géniez P., Lukanov S., Crochet P.-A., Dubey S., Perrin N., Litvinchuk S.N., Denoël M. 2019. Phylogeography of a cryptic speciation continuum in Eurasian spadefoot toads (*Pelobates*). *Molecular Ecology*. 28:3257–3270.
- Duminil J., Fineschi S., Hampe A., Jordano P., Salvini D., Vendramin G.G., Petit R.J. 2007. Can Population Genetic Structure Be Predicted from Life-History Traits? *The American Naturalist*. 169:662–672.
- Dynesius M., Jansson R. 2014. PERSISTENCE OF WITHIN-SPECIES LINEAGES: A NEGLECTED CONTROL OF SPECIATION RATES. *Evolution*. 68:923–934.
- Eaton D.A.R., Overcast I. 2020. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*. 36:2592–2594.
- Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE*. 6:e19379.
- Ferrier S., Manion G., Elith J., Richardson K. 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*. 13:252–264.
- Fick S.E., Hijmans R.J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*. 37:4302–4315.

- Fitzpatrick S.W., Brasileiro C.A., Haddad C.F.B., Zamudio K.R. 2009. Geographical variation in genetic structure of an Atlantic Coastal Forest frog reveals regional differences in habitat stability. *Molecular Ecology*. 18:2877–2896.
- Fitzpatrick S.W., Gerberich J.C., Kronenberger J.A., Angeloni L.M., Funk W.C. 2015. Locally adapted traits maintained in the face of high gene flow. *Ecology Letters*. 18:37–47.
- Forester B.R., Lasky J.R., Wagner H.H., Urban D.L. 2018. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*. 27:2215–2233.
- Freedman A.H., Harrigan R.J., Zhen Y., Hamilton A.M., Smith T.B. 2023. Evidence for ecotone speciation across an African rainforest–savanna gradient. *Molecular Ecology*. 32:2287–2300.
- Frei D., De-Kayne R., Selz O.M., Seehausen O., Feulner P.G.D. 2022. Genomic variation from an extinct species is retained in the extant radiation following speciation reversal. *Nat Ecol Evol*. 6:461–468.
- Frichot E., Francois O. 2015. LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*. 6:925–929.
- Frichot E., Mathieu F., Trouillon T., Bouchard G., Francois O. 2014. Fast and Efficient Estimation of Individual Ancestry Coefficients. *Genetics*. 196:973–983.
- Garnier S., Alibert P., Audiot P., Prieur B., Rasplus J.-Y. 2004. Isolation by distance and sharp discontinuities in gene frequencies: implications for the phylogeography of an alpine insect species, *Carabus solieri*. *Molecular Ecology*. 13:1883–1897.
- Graham S.P., Beamer D., Lamb T. 2012. Good news at last: conservation status of the seepage salamander (*Desmognathus aeneus*). *Herpetological Conservation and Biology*. 7:339–348.
- Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLoS Genet*. 5:e1000695.
- Harrison J.R. 1992. *Desmognathus aeneus*. *Catalogue of American Amphibians and Reptiles*. 534:1–4.
- Hartl D.L., Clark A.G. 2007. *Principles of population genetics*. Sunderland, Mass: Sinauer Associates.
- Harvey M.G., Seeholzer G.F., Smith B.T., Rabosky D.L., Cuervo A.M., Brumfield R.T. 2017. Positive association between population genetic differentiation and speciation rates in New World birds. *Proceedings of the National Academy of Sciences*. 114:6328–6333.
- Hey J., Pinho C. 2012. Population Genetics and Objectivity in Species Diagnosis. *Evolution*. 66:1413–1429.
- Hijmans R.J., Phillips S., Leathwick J., Elith J., Hijmans M.R.J. 2017. Package ‘dismo.’ *Circles*. 9:1–68.
- Huang J.-P. 2020. Is population subdivision different from speciation? From phylogeography to species delimitation. *Ecology and Evolution*. 10:6890–6896.
- Huang J.-P., Knowles L.L. 2016. The Species versus Subspecies Conundrum: Quantitative Delimitation from Integrating Multiple Data Types within a Single Bayesian Approach in Hercules Beetles. *Systematic Biology*. 65:685–699.
- 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.
- Ilut D.C., Nydam M.L., Hare M.P. 2014. Defining Loci in Restriction-Based Reduced Representation Genomic Data from Nonmodel Species: Sources of Bias and Diagnostics for Optimal Clustering. *BioMed Research International*. 2014:1–9.

- Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species Delimitation with Gene Flow. *Systematic Biology*. 66:799–812.
- Janzen D.H. 1967. Why Mountain Passes are Higher in the Tropics. *The American Naturalist*. 101:233–249.
- Jiao X., Flouri T., Rannala B., Yang Z. 2020. The Impact of Cross-Species Gene Flow on Species Tree Estimation. *Syst Biol*. 69:830–847.
- Jiao X., Yang Z. 2021. Defining Species When There is Gene Flow. *Systematic Biology*. 70:108–119.
- Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24:1403–1405.
- Jombart T., Devillard S., Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet*. 11:94.
- Jouganous J., Long W., Ragsdale A.P., Gravel S. 2017. Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. *Genetics*. 206:1549–1567.
- Kearns A.M., Campana M.G., Slikas B., Berry L., Saitoh T., Cibois A., Fleischer R.C. 2022. Conservation genomics and systematics of a near-extinct island radiation. *Molecular Ecology*. 31:1995–2012.
- Kisel Y., Barraclough T.G., Donohue A.E.K., McPeck E.M.A. 2010. Speciation Has a Spatial Scale That Depends on Levels of Gene Flow. *The American Naturalist*. 175:316–334.
- Kremer A., Ronce O., Robledo-Arnuncio J.J., Guillaume F., Bohrer G., Nathan R., Bridle J.R., Gómkiewicz R., Klein E.K., Ritland K., Kuparinen A., Gerber S., Schueler S. 2012. Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*. 15:378–392.
- Ledbetter N.M., Bonett R.M. 2019. Terrestriality constrains salamander limb diversification: Implications for the evolution of pentadactyly. *Journal of Evolutionary Biology*. 32:642–652.
- Leffler E.M., Bullaughey K., Matute D.R., Meyer W.K., Segurel L., Venkat A., Andolfatto P., Przeworski M. 2012. Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLOS Biology*. 10:e1001388.
- Linck E., Battey C.J. 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources*. 19:639–647.
- Liu C., Newell G., White M. 2015. On the selection of thresholds for predicting species occurrence with presence-only data. *Ecol Evol*. 6:337–348.
- MacLachlan G., Hurlburt J., Suarez M., Wong K.L., Burke W., Lewis T., Gallo A., Flidr J., Gabiam R., Nicholas J., Ensor B. 2020. Building a Shared Resource HPC Center Across University Schools and Institutes: A Case Study. *arXiv:2003.13629 [cs]*.
- Maier P.A., Vandergast A.G., Ostoja S.M., Aguilar A., Bohonak A.J. 2019. Pleistocene glacial cycles drove lineage diversification and fusion in the Yosemite toad (*Anaxyrus canorus*). *Evolution*. 73:2476–2496.
- McCartney-Melstad E., Gidiş M., Shaffer H.B. 2019. An empirical pipeline for choosing the optimal clustering threshold in RADseq studies. *Mol Ecol Resour*. 19:1195–1204.
- Mills L.S., Allendorf F.W. 1996. The One-Migrant-per-Generation Rule in Conservation and Management. *Conservation Biology*. 10:1509–1518.
- Miranda R.B. de, Brandão R.A., O'Connell K.A., Colli G.R., Tonini J.F.R., Pyron R.A. 2023. Multilocus environmental adaptation and population structure in populations of the Cerrado gecko *Gymnodactylus amarali* (Sauria, Phyllodactylidae) from Serra da Mesa Hydroelectric Plant, Central Brazil. *Frontiers in Ecology and Evolution*. in press.

- Mokany K., Ware C., Woolley S.N.C., Ferrier S., Fitzpatrick M.C. 2022. A working guide to harnessing generalized dissimilarity modelling for biodiversity analysis and conservation assessment. *Global Ecology and Biogeography*. 31:802–821.
- Moreno-Contreras I., Llanes-Quevedo A., Sánchez-González L.A., Arizmendi M.D.C., Navarro-Sigüenza A.G. 2023. Isolation by resistance explains genetic diversity in the Arremon brushfinches of northern Mesoamerica. *Mol Ecol*. 32:3450–3470.
- Muñoz M.M., Bodensteiner B.L. 2019. Janzen’s Hypothesis Meets the Bogert Effect: Connecting Climate Variation, Thermoregulatory Behavior, and Rates of Physiological Evolution. *Integrative Organismal Biology*. 1:oby002.
- Nali R.C., Becker C.G., Zamudio K.R., Prado C.P.A. 2020. Topography, more than land cover, explains genetic diversity in a Neotropical savanna tree frog. *Diversity and Distributions*. 26:1798–1812.
- Nosil P., Egan S.P., Funk D.J. 2008. Heterogeneous Genomic Differentiation between Walking-Stick Ecotypes: “Isolation by Adaptation” and Multiple Roles for Divergent Selection. *Evolution*. 62:316–336.
- Nosil P., Harmon L.J., Seehausen O. 2009. Ecological explanations for (incomplete) speciation. *Trends Ecol Evol*. 24:145–156.
- Noskova E., Abramov N., Iliutkin S., Sidorin A., Dobrynin P., Ulyantsev V. 2023. GADMA2: more efficient and flexible demographic inference from genetic data. :2022.06.14.496083.
- Noskova E., Ulyantsev V., Koepfli K.-P., O’Brien S.J., Dobrynin P. 2020. GADMA: Genetic algorithm for inferring demographic history of multiple populations from allele frequency spectrum data. *GigaScience*. 9:giaa005.
- O’Connell K.A., Prates I., Scheinberg L.A., Mulder K.P., Bell R.C. 2021. Speciation and secondary contact in a fossorial island endemic, the São Tomé caecilian. *Molecular Ecology*. 30:2859–2871.
- Omernik J.M., Griffith G.E. 2014. Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. *Environmental Management*. 54:1249–1266.
- Onn C., Grismer L. 2021. A standardized and statistically defensible framework for quantitative morphological analyses in taxonomic studies. *Zootaxa*. 5023:293–300.
- 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*. 22:5983–5999.
- Palumbi S.R., Grabowsky G., Duda T., Geyer L., Tachino N. 1997. SPECIATION AND POPULATION GENETIC STRUCTURE IN TROPICAL PACIFIC SEA URCHINS. *Evolution*. 51:1506–1517.
- Pardo-Diaz C., Salazar C., Baxter S.W., Merot C., Figueiredo-Ready W., Joron M., McMillan W.O., Jiggins C.D. 2012. Adaptive Introgression across Species Boundaries in *Heliconius* Butterflies. *PLOS Genetics*. 8:e1002752.
- Petkova D., Novembre J., Stephens M. 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nat Genet*. 48:94–100.
- Postma E., van Noordwijk A.J. 2005. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*. 433:65–68.
- Prates I., Hutchinson M.N., Singhal S., Moritz C., Rabosky D.L. 2023. Notes from the taxonomic disaster zone: Evolutionary drivers of intractable species boundaries in an Australian lizard clade (Scincidae: *Ctenotus*). *Molecular Ecology*.mec.17074.

- Pyron R.A., Beamer D.A. 2023. Systematic revision of the Spotted and Northern Dusky Salamanders (Plethodontidae: *Desmognathus conanti* and *D. fuscus*), with six new species from the eastern United States. *Zootaxa*. 5311:451–504.
- Pyron R.A., O’Connell K.A., Duncan S.C., Burbrink F.T., Beamer D.A. 2023. Speciation Hypotheses from Phylogeographic Delimitation Yield an Integrative Taxonomy for Seal Salamanders ( *Desmognathus monticola* ). *Systematic Biology*. 72:179–197.
- Pyron R.A., O’Connell K.A., Lemmon E.M., Lemmon A.R., Beamer D.A. 2020. Phylogenomic data reveal recombination and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*). *Molecular Phylogenetics and Evolution*. 146:106751.
- Pyron R.A., O’Connell K.A., Lemmon E.M., Lemmon A.R., Beamer D.A. 2022. Candidate-species delimitation in *Desmognathus* salamanders reveals gene flow across lineage boundaries, confounding phylogenetic estimation and clarifying hybrid zones. *Ecology and Evolution*. 12:e8574.
- Rissler L.J., Hijmans R.J., Graham C.H., Moritz C., Wake D.B. 2006. Phylogeographic Lineages and Species Comparisons in Conservation Analyses: A Case Study of California Herpetofauna. *The American Naturalist*. 167:655–666.
- Rosenblum E.B., Sarver B.A.J., Brown J.W., Des Roches S., Hardwick K.M., Hether T.D., Eastman J.M., Pennell M.W., Harmon L.J. 2012. Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. *Evol Biol*. 39:255–261.
- Rousset F. 1997. Genetic Differentiation and Estimation of Gene Flow from F-Statistics Under Isolation by Distance. *Genetics*. 145:1219–1228.
- Roux C., Fraisse C., Romiguier J., Anciaux Y., Galtier N., Bierne N. 2016. Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLOS Biology*. 14:e2000234.
- Schild D.R., Perry B.W., Adams R.H., Card D.C., Jezkova T., Pasquesi G.I.M., Nikolakis Z.L., Row K., Meik J.M., Smith C.F., Mackessy S.P., Castoe T.A. 2019. Allopatric divergence and secondary contact with gene flow: a recurring theme in rattlesnake speciation. *Biological Journal of the Linnean Society*. 128:149–169.
- Seeholzer G.F., Brumfield R.T. 2018. Isolation by distance, not incipient ecological speciation, explains genetic differentiation in an Andean songbird (Aves: Furnariidae: *Cranioleuca antisimensis*, Line-cheeked Spine-tail) despite near threefold body size change across an environmental gradient. *Mol Ecol*. 27:279–296.
- Sexton J.P., Hangartner S.B., Hoffmann A.A. 2014. Genetic Isolation by Environment or Distance: Which Pattern of Gene Flow Is Most Common? *Evolution*. 68:1–15.
- Singhal S., Colli G.R., Grundler M.R., Costa G.C., Prates I., Rabosky D.L. 2022. No link between population isolation and speciation rate in squamate reptiles. *PNAS*. 119.
- Singhal S., Hoskin C.J., Couper P., Potter S., Moritz C. 2018a. A Framework for Resolving Cryptic Species: A Case Study from the Lizards of the Australian Wet Tropics. *Systematic Biology*. 67:1061–1075.
- Singhal S., Huang H., Grundler M.R., Marchan-Rivadeneira M.R., Holmes I., Title P.O., Donnellan S.C., Rabosky D.L. 2018b. Does Population Structure Predict the Rate of Speciation? A Comparative Test across Australia’s Most Diverse Vertebrate Radiation. *The American Naturalist*. 192:432–447.
- Singhal S., Moritz C. 2013. Reproductive isolation between phylogeographic lineages scales with divergence. *Proc Biol Sci*. 280:20132246.
- Slatkin M. 1987. Gene Flow and the Geographic Structure of Natural Populations. *Science*. 236:787–792.
- Slatkin M. 1993. Isolation by Distance in Equilibrium and Non-Equilibrium Populations. *Evolution*. 47:264–279.

- Smadja C.M., Butlin R.K. 2011. A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*. 20:5123–5140.
- Smith M.L., Carstens B.C. 2020. Process-based species delimitation leads to identification of more biologically relevant species. *Evolution*. 74:216–229.
- Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006. Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology*. 15:4261–4293.
- Stankowski S., Ravinet M. 2021. Defining the speciation continuum. *Evolution*. 75:1256–1273.
- van Strien M.J., Holderegger R., Van Heck H.J. 2015. Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*. 114:27–37.
- Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. *Proc Natl Acad Sci USA*. 114:1607–1612.
- Title P.O., Bemmels J.B. 2018. ENVIREM: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography*. 41:291–307.
- Twyford A.D., Wong E.L.Y., Friedman J. 2020. Multi-level patterns of genetic structure and isolation by distance in the widespread plant *Mimulus guttatus*. *Heredity (Edinb)*. 125:227–239.
- Valette T., Leitwein M., Lascaux J.-M., Desmarais E., Berrebi P., Guinand B. 2023. Redundancy analysis, genome-wide association studies and the pigmentation of brown trout (*Salmo trutta* L.). *J Fish Biol*. 102:96–118.
- Velo-Anton G., Parra J.L., Parra-Olea G., Zamudio K.R. 2013. Tracking climate change in a dispersal-limited species: reduced spatial and genetic connectivity in a montane salamander. *Molecular Ecology*. 22:3261–3278.
- Wang I.J., Bradburd G.S. 2014. Isolation by environment. *Mol Ecol*. 23:5649–5662.
- Weir B.S., Cockerham C.C. 1984. Estimating F-Statistics for the Analysis of Population Structure. *Evolution*. 38:1358–1370.
- Wishingrad V., Thomson R.C. 2023. Temperate Zone Isolation by Climate: An Extension of Janzen’s 1967 Hypothesis. *The American Naturalist*. 201:302–314.
- Wright S. 1931. Evolution in Mendelian Populations. *Genetics*. 16:97–159.
- Yang L., Kong H., Huang J.-P., Kang M. 2019. Different species or genetically divergent populations? Integrative species delimitation of the *Primulina hochiensis* complex from isolated karst habitats. *Molecular Phylogenetics and Evolution*. 132:219–231.
- Zamudio K.R., Bell R.C., Mason N.A. 2016. Phenotypes in phylogeography: Species’ traits, environmental variation, and vertebrate diversification. *Proc Natl Acad Sci USA*. 113:8041–8048.
- Zbinden Z.D., Douglas M.R., Chafin T.K., Douglas M.E. 2022. Riverscape community genomics: A comparative analytical approach to identify common drivers of spatial structure. *Mol Ecol*.
- Zhang M., Suren H., Holliday J.A. 2019. Phenotypic and Genomic Local Adaptation across Latitude and Altitude in *Populus trichocarpa*. *Genome Biology and Evolution*. 11:2256–2272.

### Data Accessibility and Benefit-Sharing

*Data Accessibility Statement:* All data are available in Zenodo 10.5281/zenodo.8213268.

*Benefit-Sharing Statement:* Benefits from this research accrue from the sharing of our data and results on public databases as described above.

### Author Contributions

All authors designed and performed research, analyzed data, and wrote the paper.

### Figure Legends

**Fig. 1.** Photograph of adult *Desmognathus aeneus* from northern Georgia (courtesy Todd W. Pierson, KSU) and map of 90 sample localities with SDM results of low, medium, and high habitat suitability based on 10 BIOCLIM, ENVIREM, and ecoregion variables (see SI text).

**Fig. 2.** Linear morphometric analysis of size-corrected phenotypic traits using LDA based on 89 individuals from 31 populations, analysis of which do not generally separate the three geographic genetic lineages. Axilla-Groin Length (AG) was the most substantial variable distinguishing lineages but was not significantly different between them.

**Fig. 3.** The best-fit three-population demographic model from GADMA estimated using the ‘moments’ engine, showing topology, divergence times, population sizes, and migration rates (a), and estimates of individual ancestry coefficients from sNMF (b, d) along with a PCA of the SNP matrix and the relationship between geographic and genetic distance from the GDM (c).

**Fig. 4.** Analysis of IBD based on all individuals, between lineages, and within lineages, showing linear (all) and LOESS (a) relationship between geographic distance (km) and linearized  $F_{st}$ .

**Figure 5.** Results from RDA showing genotype-environment and genotype-phenotype association by sites (a, b) and SNPs (c, d). The 18 SNPs shared in both are outlined in red.





