




2023

PLACING THE EVOLUTIONARY HISTORY OF *DESMOGNATHUS* SALAMANDERS IN CONTEXT: A PHYLOGEOGRAPHIC APPROACH

Kara Jones

University of Kentucky, happylittleclouds@rocketmail.com

Author ORCID Identifier:

 <https://orcid.org/0000-0002-8168-0815>

Digital Object Identifier: <https://doi.org/10.13023/etd.2023.083>

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

Recommended Citation

Jones, Kara, "PLACING THE EVOLUTIONARY HISTORY OF *DESMOGNATHUS* SALAMANDERS IN CONTEXT: A PHYLOGEOGRAPHIC APPROACH" (2023). *Theses and Dissertations--Biology*. 92.
https://uknowledge.uky.edu/biology_etds/92

This Doctoral Dissertation is brought to you for free and open access by the Biology at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Biology by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Kara Jones, Student

Dr. Jeramiah J. Smith, Major Professor

Jessica C. Santollo, Director of Graduate Studies

PLACING THE EVOLUTIONARY HISTORY OF *DESMOGNATHUS*
SALAMANDERS IN CONTEXT: A PHYLOGEOGRAPHIC APPROACH

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Arts and Sciences
at the University of Kentucky

By
Kara Suzanne Jones
Lexington, Kentucky
Director: Dr. Jeramiah J. Smith, Professor of Biology
Lexington, Kentucky
2023

Copyright © Kara Suzanne Jones 2023
<https://orcid.org/0000-0002-8168-0815>

ABSTRACT OF DISSERTATION

PLACING THE EVOLUTIONARY HISTORY OF DESMOGNATHUS SALAMANDERS IN CONTEXT: A PHYLOGEOGRAPHIC APPROACH

Patterns of genetic variation do not arise in a vacuum but are instead shaped by the interplay between evolutionary forces and ecological constraints. Here, I use a phylogeographic approach to examine the role that ecology played in lineage divergence in the *Desmognathus quadramaculatus* species complex (Family: Plethodontidae), which consists of three nominal species: *D. quadramaculatus*, *D. marmoratus*, and *D. folkertsi*. Previous phylogenetic studies have shown that individuals from these species do not form clades based on phenotype. My approach to reconciling phylogenetic discordance was two-fold, using (1) genome-wide markers to provide insight into the relationships among lineages and (2) geographic and climate data to provide context for patterns of genetic diversity.

First, I obtained genome-wide nuclear markers using double-digest restriction-site associated DNA sequencing (ddRAD) to examine whether two morphologically divergent species, *D. marmoratus* and *D. quadramaculatus*, represent independently evolving lineages. Phylogenetic, population structure, and model testing analyses all confirmed that *D. marmoratus* and *D. quadramaculatus* do not group based on phenotype. Instead, I found that there were two cryptic genetic lineages (Nantahala and Pisgah) that each contained both phenotypes. Additionally, ecological niche modeling showed that the two genetic lineages primarily occupy geographic areas with significantly different climates, suggesting that climate may have played a role in divergence.

Next, I assembled loci from publicly available sequencing data using a draft transcriptome of *Desmognathus fuscus* as a reference to assess the three nominal species in the quadramaculatus species complex across their entire range. I used phylogenetic and population structure analyses, alongside haplowebs and conspecificity matrices, to determine if the loci supported the hypothesis that the phenotypes represent multiple independently evolving lineages within the broader genetic clades found in the previous chapter. I found that the loci were not informative enough to determine whether the phenotypes had a genetic basis in Pisgah, but did support genetic divergence between phenotypes in Nantahala.

Finally, I used ecological niche models (ENMs) and resistance modeling to place the genetic results and phenotypic diversity within the context of time and space. I found that though the quadramaculatus and marmoratus phenotypes were nearly indistinguishable in niche space in the present day, they were projected to occupy different geographic areas in the past and future. The southern portion of the study area

had areas of high habitat suitability from the Last Glacial Maximum (~22 kya) to the present, which aligns with the higher genetic divergence between groups in Nantahala. Anthropogenic land use changes reduced habitat availability but likely did not drive genetic divergence in the past, and may be of more consequence to genetic diversity than climate change over the next 50 years.

Like many taxa that underwent adaptive radiations, the evolutionary history of *Desmognathus* has been obfuscated by high rates of within-species phenotypic diversity and shared morphology between distantly related lineages. My findings emphasize the importance of interrogating complex patterns of genetic variation within the context of the dynamic, heterogeneous landscapes in which they arise.

KEYWORDS: Phylogeography, Phylogenetics, Population Genetics, Ecological Niche Modeling, Ecological Speciation, *Desmognathus*

Kara Suzanne Jones

03/30/2023

Date

PLACING THE EVOLUTIONARY HISTORY OF DESMOGNATHUS
SALAMANDERS IN CONTEXT: A PHYLOGEOGRAPHIC APPROACH

By
Kara Suzanne Jones

Jeramiah J. Smith

Director of Dissertation

Jessica C. Santollo

Director of Graduate Studies

03/30/2023

Date

TABLE OF CONTENTS

| | |
|---|------|
| LIST OF TABLES | viii |
| LIST OF FIGURES | ix |
| CHAPTER 1. INTRODUCTION | 1 |
| 1.1 Ecological opportunity and divergence | 1 |
| 1.2 Phylogeography and climate change | 2 |
| 1.3 Distribution and life history of the <i>Desmognathus quadramaculatus</i> species complex | 4 |
| 1.4 Overview of dissertation chapters | 6 |
| CHAPTER 2. GENOMIC DATA REJECT THE HYPOTHESIS OF SYMPATRIC ECOLOGICAL SPECIATION IN A CLADE OF <i>DESMOGNATHUS</i> SALAMANDERS (FAMILY: PLETHODONTIDAE) | 7 |
| 2.1 Abstract | 7 |
| 2.2 Introduction | 8 |
| 2.3 Methods | 11 |
| 2.3.1 Tissue sampling and sequencing | 11 |
| 2.3.2 Population structure | 14 |
| 2.3.3 Phylogenetics | 15 |
| 2.3.4 Testing hypotheses of lineage divergence | 16 |
| 2.3.5 Gene flow among lineages | 17 |
| 2.3.6 Climate niche modeling | 18 |
| 2.4 Results | 21 |
| 2.4.1 Population structure | 22 |
| 2.4.2 Phylogenetics | 23 |
| 2.4.3 Testing hypotheses of lineage divergence | 24 |
| 2.4.4 Gene flow among lineages | 24 |
| 2.4.5 Climate niche modeling | 25 |
| 2.5 Discussion | 25 |
| 2.5.1 Possible mechanisms underlying morphological differentiation | 26 |
| 2.5.2 Geographic split corresponds to lineage divergence | 28 |
| 2.5.3 Broader phylogenetic perspective | 31 |
| CHAPTER 3. MITOCHONDRIAL AND NUCLEAR DATA SUPPORT LINEAGE DIVERGENCE WITHIN THE <i>DESMOGNATHUS QUADRAMACULATUS</i> SPECIES COMPLEX IN NANTAHALA BUT NOT PISGAH | 40 |
| 3.1 Abstract | 40 |
| 3.2 Introduction | 41 |
| 3.3 Methods | 43 |
| 3.3.1 Nuclear data preparation | 43 |
| 3.3.1.1 Raw sequence processing | 43 |
| 3.3.1.2 Read mapping, variant calling, and candidate locus assembly .. | 44 |

| | | |
|---|--|-----|
| 3.3.1.3 | Final locus selection and data set construction | 45 |
| 3.3.1.4 | Comparison between desmo900-all and Pyron22 loci..... | 45 |
| 3.3.2 | Mitochondrial data preparation | 46 |
| 3.3.3 | Genetic diversity, isolation by distance, and population structure ... | 46 |
| 3.3.4 | Population structure | 47 |
| 3.3.5 | Phylogenetic analyses | 48 |
| 3.3.6 | Haplotype networks, haplowebs, and conspecificity matrices | 49 |
| 3.4 | Results..... | 50 |
| 3.4.1 | Read mapping, variant calling, and locus assembly | 50 |
| 3.4.2 | Genetic diversity, isolation by distance, and population structure ... | 51 |
| 3.4.3 | Phylogenetic analyses | 52 |
| 3.4.4 | Haplotype networks, haplowebs, and conspecificity matrices | 53 |
| 3.5 | Discussion | 55 |
| 3.5.1 | Genetic structure in Pisgah is likely driven by drift | 55 |
| 3.5.2 | Some of the candidate species do not form well-supported clades .. | 58 |
| 3.5.3 | Haplotype networks are consistent with concordant mitonuclear divergence between Pisgah and Nantahala | 61 |
| 3.5.4 | Revisiting the ecological role of phenotypes in divergence | 63 |
| CHAPTER 4. ECOLOGICAL NICHE MODELING PROVIDES INSIGHT INTO THE DISTRIBUTION OF GENETIC AND PHENOTYPIC DIVERGENCE | | 87 |
| 4.1 | Abstract | 87 |
| 4.2 | Introduction | 88 |
| 4.3 | Methods..... | 92 |
| 4.3.1 | Study area and taxa | 92 |
| 4.3.2 | Occurrence and background data..... | 93 |
| 4.3.3 | Environmental data | 94 |
| 4.3.4 | Model calibration and evaluation..... | 96 |
| 4.3.5 | Prediction of past and future habitat suitability | 98 |
| 4.3.6 | Changes in land use between 1650-2020..... | 99 |
| 4.4 | Results..... | 101 |
| 4.4.1 | Ecological niche models | 101 |
| 4.4.2 | Past habitat suitability predictions (LGM to present) | 102 |
| 4.4.3 | Future habitat suitability predictions | 103 |
| 4.4.4 | Changes in land use between 1650-2020..... | 104 |
| 4.5 | Discussion | 105 |
| 4.5.1 | Impact of variable selection on ENMs..... | 105 |
| 4.5.2 | Historical habitat suitability | 107 |
| 4.5.3 | Effects of anthropogenic land use on population structure..... | 109 |
| 4.5.4 | Effects of anthropogenic change on habitat suitability | 111 |
| CHAPTER 5. CONCLUSION..... | | 137 |
| 5.1 | Evidence for divergence within the Nantahala lineage..... | 138 |
| 5.2 | Evidence for divergence within the Pisgah lineage | 140 |

| | |
|---|-----|
| APPENDICES | 144 |
| APPENDIX A. SUPPLEMENTARY INFORMATION FROM JONES AND WEISROCK (2018) | 145 |
| APPENDIX B. INDIVIDUALS USED IN ANALYSES | 160 |
| REFERENCES | 165 |
| VITA | 188 |

LIST OF TABLES

| | |
|--|-----|
| Table 2.1. Results from replicate Bayesian Phylogeography & Phylogenetics..... | 33 |
| Table 3.1. Genotyping rate and data set sizes for nuclear loci used in this study..... | 66 |
| Table 3.2. Summary of pairwise genetic distance (D_{PS}) for individuals within groups and between groups. | 67 |
| Table 3.3. Number of private, singletons, and diagnostic alleles found within groups.... | 68 |
| Table 3.4. Gene concordance factors and site concordance factors. | 69 |
| Table 4.1. Number of occurrence points and background points used to build ecological niche models (ENMs). | 115 |
| Table 4.2. Bioclimatic variables used in baseline models. | 116 |
| Table 4.3. Reclassification of land use and land cover (LULC) rasters used in ENMs and habitat stability analyses. | 117 |
| Table 4.4. Results for “best” models for each combination of variables and backgrounds tested. | 118 |
| Table 4.5. Variable permutation importance for CHELSA models with GR background, with and without LULC. | 119 |
| Table 4.6. Identity tests for best CHELSA+LULC ENMs. | 120 |
| Table A1. Number of <i>D. quadramaculatus</i> , <i>D. marmoratus</i> , and <i>D. carolinensis</i> at each sample site..... | 145 |
| Table A2. Number of loci removed by major filtering parameters in iPyrad for the ingroup and outgroup data sets. | 146 |
| Table A3. Parameters for the best models for each background size from ENMeval.... | 147 |
| Table A4. Results of ecological niche models for Nantahala and Pisgah. | 148 |

LIST OF FIGURES

| | |
|--|-----|
| Figure 2.1. Sampling locations and phenotypic differences. | 35 |
| Figure 2.2. Models for Bayesian Phylogeography & Phylogenetics. | 36 |
| Figure 2.3. Results from Discriminant Analysis of Principle Components (DAPC). | 38 |
| Figure 2.4. MaxEnt environmental niche models. | 39 |
| Figure 3.1. Sampling map of individuals used in this study. | 70 |
| Figure 3.2. Euclidean genetic distance versus geographic distance for all individuals. | 72 |
| Figure 3.3. Principal component analysis. | 73 |
| Figure 3.4. Population structure results from sNMF for Pisgah. | 74 |
| Figure 3.5. Population structure results from sNMF for Nantahala. | 75 |
| Figure 3.6. Six gene properties used to assess phylogenetic usefulness in genesortR for the 221 loci used in the desmo900-all phylogeny. | 76 |
| Figure 3.7. Maximum likelihood phylogenetic tree built using a partitioned concatenated matrix of desmo900-all loci. | 77 |
| Figure 3.8. TCS haplotype network for COI. | 80 |
| Figure 3.9. TCS haplotype network for CYTB. | 82 |
| Figure 3.10. Example haplowebs for five nuclear loci. | 84 |
| Figure 3.11. Consppecificity matrix analysis using the sum of conspecific sites. | 85 |
| Figure 3.12. Results of conspecificity matrix analysis using the sum of conspecific loci minus the sum of heterospecific loci. | 86 |
| Figure 4.1. Shaded relief map of study area and occurrence locations where individuals of the <i>Desmognathus quadramaculatus</i> species complex have been found within the last 30 years (1992-2022). | 121 |
| Figure 4.2. Comparison of two bioclimatic variables for Worldclim 2.1 and CHELSA 1.4. | 122 |
| Figure 4.3. Multivariate Environmental Similarity Surfaces (MESS). | 123 |
| Figure 4.4. Habitat suitability maps for pheotypes. | 124 |
| Figure 4.5. Habitat suitability maps for genetic clades. | 125 |
| Figure 4.6. Visualization of results for identity by environment tests for CHELSA+LULC. | 126 |
| Figure 4.7. Habitat suitability maps for phenotypes projected to the Last Glacial Maximum (LGM). | 128 |
| Figure 4.8. Habitat suitability maps for phenotypes projected to 2050 and 2070. | 130 |
| Figure 4.9. Percent forest cover in study region from 1650-2000. | 131 |
| Figure 4.10. Difference in percent forest cover between 1650 and 2000. | 132 |
| Figure 4.11. Forest stand age found at 5 km grid background points and occurrence points. | 133 |
| Figure 4.12. Habitat stability from 1650-2000. | 134 |
| Figure 4.13. Land use and land cover for 1650 and 2000. | 135 |
| Figure 4.14. Correlation between genetic distance and resistance distance between 1650 and 2000. | 136 |
| Figure A1. Graph of cross-validation error for 10 runs of Admixture for $K = 1-8$ showing an optimal K of 2. | 149 |
| Figure A2. Map of sampling sites and Admixture results for $K = 3$ | 150 |

| | |
|---|-----|
| Figure A3. Plots of assignment probabilities from Discriminant Analysis of Principle Components (DAPC). | 151 |
| Figure A4. Unrooted SVDquartets tree. | 152 |
| Figure A5. TreeMix plot showing relationships between groups based on the four deepest splits recovered in the BEAST ultrametric tree. | 154 |
| Figure A6. Map of sampling points used for ecological niche models. | 155 |
| Figure A7. Results of symmetrical background test (A & B) and identity test (C & D) 156 | |
| Figure A8. Blob range-break test results for Nantahala versus Pisgah..... | 157 |
| Figure A9. Map showing relationship between major genetic divisions within <i>D. quadramaculatus</i> and river basins. | 158 |
| Figure A10. Map of annual precipitation (BIO12) across southern Appalachia. | 159 |

CHAPTER 1. INTRODUCTION

1.1 Ecological opportunity and divergence

Lineage divergence is often driven by the availability of new ecological opportunities (Ceccarelli et al. 2016; De-Silva et al. 2016; Pérez-Escobar et al. 2017; Condamine et al. 2018). For example, climate shifts can fracture niches into narrow or heterogeneous patches, allowing specialists to diverge from generalist ancestors (Martínez-Cabrera and Peres-Neto 2013; Ezard and Purvis 2016). Increased topographic complexity created during mountain uplift can create a range of narrow niches composed of varying horizontal microclimates (e.g. high precipitation versus rain shadow) and vertical microhabitats (e.g. forests versus grassy balds), leading to a higher number of species in montane landscapes (Condamine et al. 2018). Thus, a dynamic landscape can facilitate lineage divergence by increasing available niche space.

The evolution of phenotypes to exploit novel niche space may involve co-opting existing adaptations for new uses (i.e., exaptation) or by differential selection on existing polymorphisms (Rabosky and Lovette 2008; Losos and Malher 2010; Berner and Salzburger 2015). New phenotypes can also be created during secondary contact (Mallet 2007; Jones et al. 2012; Lamichhaney et al. 2015), which can provide an influx of new genetic variation for selection to act on (Seehausen 2004; Hedrick 2013; Arnold and Kunte 2017). Conversely, lineage divergence can result in multiple species converging on parallel phenotypes, leading to cryptic speciation (Bickford et al. 2007; Vodă et al. 2015; Espíndola et al. 2016).

Lineages capable of exploiting a wide range of ecological opportunities are considered to have high diversification potential (Wellborn and Langerhans 2015). Diversification potential can be broadly defined as the ability of a lineage to diversify into a variety of niches and then convert those divergent phenotypes into species. This transition from divergent phenotypes to “good” species can be generalized to three stages: (1) functional divergence, when individual genes diverge without accompanying genomic divergence, (2) ecological divergence, when genomic islands of divergence form, and (3) lineage divergence, when genome-wide divergence forms (Wu 2001; Seehausen et al. 2014). Plasticity may play a role in the initial stages of divergence, but lineage divergence ultimately results in phenotypes that are canalized and heritable (Pfennig et al. 2010; Thibert-Plante and Hendry 2011). Furthermore, there is no guarantee that phenotypic divergence will result in lineage divergence (DeWitt and Scheiner 2004; De Jong 2005). Hence, phenotypes can provide misleading clues when studying lineage divergence.

1.2 Phylogeography and climate change

Phylogeography aims to provide a biogeographical context for phylogenetics and population genetics (Avice 1987). Phylogenetics and population genetics both examine how evolutionary processes have affected the development and maintenance of species, with phylogenetics focusing on the distribution of genetic diversity among lineages, and population genetics focusing within lineages. By looking at how this genetic diversity is distributed across the landscape, phylogeography can provide insight into the ecological processes that have led to lineage divergence and population structure (Avice 2000;

Hickerson et al. 2009). While there are disagreements about what constitutes the specific purview of phylogeography rather than other closely related fields (e.g., historical biogeography [Wiens 2012]; landscape genetics [Rissler 2016]), in practice the term is used to describe studies at a variety of focal breadths (i.e., individuals, populations, lineages, and communities) and across a range of spatial (i.e., local, regional, and global) and temporal (i.e., contemporary and historical) scales (Beheregaray 2008; Avise 2009).

Phylogeography has been used extensively to study the effects of past climate change on lineages. The climate oscillations that defined the Quaternary Period (~2.6 mya to present) resulted in high rates of habitat instability, driving the movement of populations across the landscape as they tracked shifting areas of suitable habitat (Hewitt 2000; Davis and Shaw 2001; Nogués-Bravo et al. 2007). Areas distant from continental glaciation are often considered more stable, providing refuges where species can “wait out” ice ages. However, within areas spared from glaciation, landscape heterogeneity may have led to pockets of stable habitat surrounded by unsuitable habitat (Schmitt and Varga 2012). These climate refuges potentially acted as islands of climate stability and storehouses of genetic diversity during glacial maxima, safeguarding the raw materials necessary for successful recolonization of newly habitable areas during interglacial periods. Thus refuges probably played a large role in shaping the genetic diversity in two ways: (1) previously allopatric populations that came to share the same refuge may have been squeezed into parapatry and sympatry, leading to opportunities for gene flow (Hewitt 2011; Thesing et al. 2016; Allen et al. 2021), and (2) populations that were split between separate refuges became increasingly isolated, leading to lineage divergence (Crespi et al. 2003; Hedin et al. 2015; Hughes and Atchison 2015; Zhang et al. 2019).

For species shaped by past climate change, repeated retraction into and expansion out of climate refuges led to periods of genetic divergence followed by secondary contact after relatively short (<100 kya) time spans (Hewitt 2000). These periods of divergence may not have lasted long enough to produce full reproductive isolation; hence retreat back into climate refuges created opportunities for previously separated populations to engage in gene flow. Refuges were therefore places where populations could rejuvenate diversity lost through drift, leading to post-glaciation populations with genomes likely reflecting complex reticulate evolutionary history.

1.3 Distribution and life history of the *Desmognathus quadramaculatus* species complex

One climate refuge whose history remains understudied is the biodiversity hotspot of southern Appalachia. Southern Appalachia is one of the most topographically complex regions in North America and houses high numbers of endemic species across a range of taxa. The salamander genus, *Desmognathus* (family: Plethodontidae), lends itself to studying the process of phenotypic and genetic diversification in southern Appalachia because it contains the most diverse array of life histories of any salamander group in the region. In this dissertation, I focus specifically on the *Desmognathus quadramaculatus* species complex—consisting of *D. quadramaculatus*, *D. marmoratus*, and *D. folkertsi*—which were originally described as separate species based on their divergent phenotypes, proposed to be the result of ecological speciation (Martof 1956; Camp et al. 2002; Wooten and Rissler 2011). However, previous studies have indicated that there may be

cryptic divergence within this lineage (Rissler and Taylor 2003; Kozak et al. 2005; Jones et al. 2006; Camp and Wooten 2016), and its evolutionary history remains uncertain.

This complex is primarily found within the Blue Ridge ecoregion in the southern Appalachian Mountains, including western North Carolina, southwestern Virginia, eastern Tennessee, northeastern Georgia, and northwestern South Carolina. The geographic distribution of the *marmoratus* and *folkertsi* phenotypes nest within the range of *quadramaculatus*, and phenotypes co-occur at some sites where their ranges overlap. The *marmoratus* phenotype is only found in the Blue Ridge ecoregion while the *quadramaculatus* phenotype extends into adjacent portions of the Piedmont to the east, Ridge and Valley in southwestern Virginia, and Central Appalachians in West Virginia. The *folkertsi* phenotype is confined to a small area of the southern Blue Ridge in northern Georgia, southwestern North Carolina, and northeastern South Carolina.

All of the phenotypes primarily inhabit fast-flowing, highly-oxygenated mountain streams in mesic forests between 300-1700 m (Petranka 2010). The aquatic larval stage occurs in the same streams where adults are found, but time to metamorphosis varies by phenotype, with *folkertsi* having a two-year larval stage and *quadramaculatus* up to a four-year larval stage (Bruce 1988; Camp et al. 2002). Adults may disperse via streams or across land and can also be found in seeps and foraging on the forest floor. As members of the family Plethodontidae, desmognathines lack lungs and respire through their skin, thus requiring a moist environment for adequate gas exchange (Ruben and Boucot 1989). Hence, all of the phenotypes in this study are highly dependent on the availability of mesic forest habitats and streams that persist long enough to ensure a successful larval stage.

1.4 Overview of dissertation chapters

The goal of this dissertation is to resolve the evolutionary history of the *D. quadramaculatus* species complex using a phylogeographic approach. In the first empirical chapter, I use genome-wide nuclear markers to identify lineage divergence across *D. quadramaculatus* and *D. marmoratus* using phylogenetic and population genetic methods. I also specifically test hypotheses regarding whether phenotypic or geographic divergence more accurately reflects the evolutionary history of the group using a Bayesian phylogenetic framework. Finally, I use ecological niche models to assess whether climate may have played a role in divergence. This chapter was previously published in *Evolution*.

In the second empirical chapter, I re-analyzed data from previous studies using a variety of methods to interrogate whether there is additional divergence within the two lineages I identified in the previous chapter. Along with phylogenetic and population genetic approaches, I added additional analyses to identify diagnostic alleles and create haplotype networks to determine whether there was evidence to support previously hypothesized candidate species.

In the final empirical chapter, I focused on providing context to the genetic findings from the previous chapters by building ecological niche models for genetic lineages and phenotypes. I projected the results into past climate scenarios to assess how suitable habitat may have changed over time, thus shaping lineage divergence. I also assessed the impact of anthropogenic changes on land use since European colonization and projected niche models onto future climate scenarios to determine how the species complex may respond to ongoing climate and land use changes.

CHAPTER 2. GENOMIC DATA REJECT THE HYPOTHESIS OF SYMPATRIC ECOLOGICAL SPECIATION IN A CLADE OF *DESMOGNATHUS* SALAMANDERS (FAMILY: PLETHODONTIDAE)

This chapter has been previously published as:

Jones, K. S., & Weisrock, D. W. (2018). Genomic data reject the hypothesis of sympatric ecological speciation in a clade of *Desmognathus* salamanders. *Evolution*, 72(11), 2378–2393. <https://doi.org/10.1111/evo.13606>

2.1 Abstract

Closely related taxa with dissimilar morphologies are often considered to have diverged via natural selection favoring different phenotypes. However, some studies have found these scenarios to be paired with limited or no genetic differentiation.

Desmognathus quadramaculatus and *D. marmoratus* are sympatric salamander species thought to represent a case of ecological speciation based on distinct morphologies, but the results of previous studies have not resolved corresponding patterns of lineage divergence. Here, we use genome-wide data to test this hypothesis of ecological speciation. Population structure analyses partitioned individuals geographically, but not morphologically, into two adjacent regions of western North Carolina: Pisgah and Nantahala. Phylogenetic analyses confirmed the nominal species are non-monophyletic and resolved deep divergence between the two geographic clusters. Model-testing overwhelmingly supported the hypothesis that lineage divergence followed geography. Finally, ecological niche modeling showed that Pisgah and Nantahala individuals occupy different climatic niches, and geographic boundaries for the two lineages correspond to differences in precipitation regimes across southern Appalachia. Overall, we reject the

previous hypothesis of ecological speciation based on microhabitat partitioning. Instead, our results suggest that there are two cryptic lineages, each containing the same pair of morphotypes.

2.2 Introduction

Species that occupy separate niches and exhibit different morphologies, particularly when distributed in sympatry, are often presumed to represent cases of ecological speciation, whereby natural selection drives the evolution of traits concomitantly with reproductive isolation (Schluter 2001; Bolnick and Fitzpatrick 2007; Schuler et al. 2016). However, even when morphological and environmental differences seem to be linked, the evolution of these differences does not always result in genome-wide divergence. Instead, morphologies associated with ecological differentiation may be tied to genomic islands of divergence (Turner et al. 2005; Guerrero and Hahn 2017; Wolf and Ellegren 2017), where genetic divergence is highly localized to key areas of the genome under selection (Twyford and Friedman 2015; Marques et al. 2016; Toews et al. 2016). In other cases there may be no consistent pattern of genetic differentiation, suggesting plasticity alone could be responsible for ecological and morphological differences, resulting in phenotypic differentiation without lineage divergence (e.g. (Dowle et al. 2014; Mason and Taylor 2015; Branch et al. 2017; Brock et al. 2017)). Thus, both ecological and morphological differentiation may act as poor proxies for genetic divergence across the speciation continuum.

Hence, a key step in elucidating whether two or more species represents a case of ecological divergence is to survey for patterns of differentiation across the genome (Seehausen et al. 2014; Wolf and Ellegren 2017). Recent studies using genome-wide

sequencing have reaffirmed support for the hypothesis of ecological speciation in some cases [e.g. Darwin's finches (Lamichhaney et al. 2015); sticklebacks (Roesti et al. 2015); pupfishes (Martin and Feinstein 2014)], but failed to do so in others [e.g. mountain chickadees (Branch et al. 2017); *Mimulus* sp. (Twyford and Friedman 2015); *Littorina* sp. (Ravinet et al. 2016)]. However, the use of genome-wide data to explore questions of ecological divergence has typically concentrated on a few groups (e.g. *Heliconius* butterflies, cichlids, sticklebacks, and monkey flowers), leaving many taxa understudied.

Morphological characters are often used in identifying putative cases of ecological speciation, but speciation may not result in diagnostic morphological characters that mark the formation of new lineages. One method of detecting these cases is to test for concordance between niche space and genetic structure at the landscape level (Rissler and Apodaca 2007; Dagnino et al. 2017). For example, ecological niche modeling can test for a significant difference in niche space usage between two putative species (Graham et al. 2004; Rosa et al. 2015; Ryan et al. 2017), thus aiding in identifying cryptic ecological species and visualizing the boundaries between them (e.g. Zhang et al. 2014; Vodă et al. 2015; Darwell and Cook 2017).

Here, we test a hypothesis of ecological speciation in the salamander genus *Desmognathus* (Caudata: Plethodontidae). Originating in southern Appalachia, this speciose genus underwent an adaptive radiation over the last 5-10 million years, resulting in members that vary widely in habitat use, body size, and life history (Wiens et al. 2003; Kozak et al. 2005, 2009; Wake et al. 2011). Two species, *D. quadramaculatus* and *D. marmoratus*, are thought to represent an example of sympatric ecological speciation. These two species overlap throughout most of their respective ranges (Figure 2.1A), and

co-occur in the same streams when in sympatry (Vences and Wake 2007). There is evidence that they micropartition habitat within streams as both larvae and adults, with adult *D. marmoratus* primarily occupying the center of streams and *D. quadramaculatus* the edges (Martof 1962; Bruce 1985). They also exhibit distinct cranial morphologies; the skull shape of *D. marmoratus*—considered an adaptation to their more aquatic lifestyle—is unique within the genus *Desmognathus* and led to them originally being placed in their own genus (Dunn 1926).

To date, multiple molecular studies—primarily based on mitochondrial DNA—have failed to find evidence for lineage divergence between *D. marmoratus* and *D. quadramaculatus*, with gene trees reconstructing a collective clade containing a non-monophyletic assemblage of haplotypes sampled from both species (Rissler and Taylor 2003; Jackson 2005; Kozak et al. 2005; Jones et al. 2006; Wooten and Rissler 2011). Consequently, species boundaries remain defined solely by morphology. Camp and Wooten (2016) suggested the phylogenetic relationships in this complex may be blurred by the presence of cryptic species and the repeated evolution of similar phenotypes. However, the paucity of genetic markers used in previous studies failed to provide definitive conclusions, leaving the relationship between the two species unclear.

While the large, highly repetitive genomes of salamanders present unique challenges (Keinath et al. 2015), these factors do not present an inherent barrier to the genomic study of ecological divergence (Weisrock et al. 2018). Thus, we tested hypotheses of ecological speciation between sympatrically distributed *D. quadramaculatus* and *D. marmoratus* using genome-wide data. To explore the relationship between the putative species in an unbiased manner, we made no assumptions as to where they fell on the speciation

continuum while investigating the following questions: (1) How are individuals structured into populations and is there evidence for genetic divergence between phenotypes? (2) Do phylogenetic and model-based perspectives support a particular hypothesis for these patterns of divergence? (3) What ecological factors may have influenced divergence in this system? We find no evidence for genome-wide divergence between the species as currently recognized. Our data reject a hypothesis that ecological speciation resulted from morphological adaptations to different aquatic habitats. Instead, we identify deep patterns of genome-wide divergence that correspond to geography, dividing the group into two independent lineages that occupy distinct climatic niches. Our results have broader implications that extend beyond this genus, as many organismal systems may harbor signatures of ecological divergence that do not correspond with their morphological differences.

2.3 Methods

2.3.1 Tissue sampling and sequencing

Collection of tissue samples took place primarily within Pisgah National Forest, Nantahala National Forest, and Great Smoky Mountains National Park (North Carolina Wildlife Resources Commission License #15-SC01018; National Park Service Scientific Research and Collecting Permit #GRSM-2015-SCI-1232). Search sites were selected based on known occurrence data obtained from the Global Biodiversity Information Facility (GBIF) for both *D. quadramaculatus* and *D. marmoratus*. A focus was placed on sampling from regions of overlap for both species, including sites where they could be collected in sympatry, and therefore we did not include regions where only *D.*

quadramaculatus is found (e.g. parts of Virginia and West Virginia). We hand-captured salamanders, sampled tail tissue, and then released the salamanders at their point of capture. Animal care and use in this study was approved by the Highlands Biological Station IACUC (HBS 15-02).

Because the *Desmognathus* genome is ~15 Gbp (Hally et al. 1986), population-level surveys of genome-wide variation in this taxon are most feasible using reduced-representation approaches. Therefore, we used double-digest restriction-site associated DNA sequencing (ddRAD) to sample loci throughout the genome (Peterson et al. 2012; Andrews and Luikart 2014). DNA was first extracted using a DNeasy Blood and Tissue Kit (Qiagen), following the manufacturer's tissue extraction protocol, and then prepared for sequencing as per Peterson et al. (2012). Briefly, extracted DNA was double-digested using equal amounts of the restriction enzymes *EcoRI* and *SphI* (New England BioLabs). Fragments from each individual were uniquely indexed and then pooled for size selection of 500 bp \pm 10% on a Pippin Prep (Sage Science). The resulting libraries were amplified, prepared for Illumina sequencing, cleaned, and sequenced with 150 bp paired-end reads on an Illumina HiSeq 2500 at the Florida State University College of Medicine and an Illumina HiSeq 4000 at the Roy J. Carver Biotechnology Center at the University of Illinois, with a sequencing effort of approximately 12 million reads per individual.

We sequenced a total of 97 field-sampled individuals from 28 localities, with an additional 17 *D. quadramaculatus* and 3 *D. marmoratus* tissue samples from 11 localities used in (Wooten and Rissler 2011). Sequence data were processed using the IPYRAD v0.6.27 pipeline (Eaton 2014) to cluster loci and call single nucleotide polymorphisms (SNPs(Eaton 2014)). Reads were first demultiplexed using individual-specific index

sequences. Individuals with fewer than one million raw reads were removed from further analyses to prevent genotyping bias resulting from variable sequencing effort among individuals. Quality filtering was applied to remove adapter sequences and fragments without cut sites or unique indices. Base calls with a Phred quality score below 20 and sites with more than two bases were converted to Ns, and each final consensus sequence allowed a maximum of four Ns. Data generated on the HiSeq 4000 produced a high proportion of low-quality base calls on the second reads; therefore, all second reads were discarded, and further processing used only the first reads. The clustering depth for assembling loci within individuals was set at a minimum of 10 reads per locus to increase the accuracy of identifying SNP sites. The clustering threshold was set at 0.85 similarity for both within- and across-individual clustering steps, which is the recommended default for IPYRAD (Eaton 2014). The maximum number of heterozygous sites shared across individuals was set at three to reduce the probability of clustering paralogs.

After processing, 59 individuals remained with data of high enough quality to use in onward analyses: 13 *D. marmoratus*, 43 *D. quadramaculatus*, and 3 *D. carolinensis* genotyped from 29 sampling locations (Figure 2.1B; Table A1). Two data sets were produced to use in most analyses: (1) an ingroup data set using only *D. quadramaculatus* and *D. marmoratus* individuals, and (2) an outgroup data set which also included two *D. carolinensis* individuals to serve as an outgroup. Due to high locus drop-out when the third *D. carolinensis* was included, we omit it from most analyses. Loci were retained if they were found in at least 85% or 75% of individuals for the ingroup and outgroup data sets, respectively. Additional data sets created for specific analyses are described below

in the respective methods sections. All genotype-based analyses used one SNP randomly selected from each locus except where noted otherwise.

2.3.2 Population structure

We used the model-based program ADMIXTURE v1.3 (Alexander et al. 2009; Alexander and Lange 2011) to test the hypothesis that *D. marmoratus* and *D. quadramaculatus* individuals would assign to different genetic clusters. We performed 10 independent replicates for $K = 1$ to 8 using the ingroup data set. We used CLUMPAK (Kopelman et al. 2015) to check for consistency of results across replicates. The optimal K was selected based on the mean error from 1000-fold cross-validation (Alexander and Lange 2011); however, because “suboptimal” K s can provide useful information about population structure, we explored multiple values of K when assessing results (Meirmans 2015).

We also used the non-parametric Discriminant Analysis of Principle Components (DAPC; Jombart et al. 2010) from the *adeigenet* package in R v3.4.0. While Admixture partitions individuals into clusters, it does not provide information about the magnitude of differentiation between those clusters. In contrast, DAPC allows us to determine the relative divergence between clusters in discriminant space. Therefore, we used DAPC to test whether individuals showed maximal differentiation within discriminant space when pre-assigned cluster membership by putative species (i.e., *D. quadramaculatus* or *D. marmoratus*) or by clusters identified in ADMIXTURE (which divided individuals by geography). We used the outgroup data set, permitting *D. carolinensis* to serve as a point of reference, to assess the magnitude of differentiation between putative species and population clusters. We used the *poppr* package to find the number of principle

components (PCs) with the highest mean assignment success according to 1000-fold cross-validation (Kamvar et al. 2015). We visualized individual assignment probabilities using the *compoplot* function to determine if individuals could be correctly assigned to their original species or population, and we generated a scatter plot to visualize how individuals related to each other within discriminant space (Jombart 2008).

We calculated unbiased pairwise F_{ST} using Hudson's estimator in the R package *PopGenome* (Bhatia et al. 2013; Pfeifer et al. 2014) for the clusters identified in ADMIXTURE at $K = 2$ (Pisgah and Nantahala); $K = 2$ clusters plus the outgroup *D. carolinensis*; and *D. marmoratus* *D. quadramaculatus* chose F_{ST} as the preferred measure of differentiation because it produces unbiased results for SNP data compared to other commonly used statistics (Meirmans and Hedrick 2011).

2.3.3 Phylogenetics

We used two approaches to evaluate the placement of individuals in a phylogenetic framework. First, we generated a coalescent-based lineage tree for all individuals using SVDQUARTETS implemented in PAUP* v4.0a157 (Swofford 2003; Chifman and Kubatko 2014). We ran two analyses, one using the ingroup data set and one using the outgroup data set with *D. carolinensis* pre-assigned as an outgroup. In both cases, we used exhaustive quartet sampling and used 1000 bootstrap replicates to determine node support.

We also constructed a rooted ultrametric tree using concatenated full-locus data from the outgroup data set (~140 bp per ddRAD locus; 462,043 total bp) using BEAST v2.2 (Bouckaert et al. 2014). We included *D. carolinensis* in the analysis, but did not pre-

assign it as an outgroup, allowing us to produce a rooted tree without prior guidance. We estimated the best-fitting substitution model for the concatenated data set as TPM1uf+I+G using BIC criteria in jModelTest v2.1.10 (Darriba et al. 2012). This was implemented in BEAST as a GTR model with equal rates for AC↔GT, AG↔CT, and AT↔CG substitutions (Kimura 1981; Posada 2008). We used a Yule pure-birth model for the tree prior with a uniform birth rate, and a relaxed log normal clock without time calibration. We performed two replicate runs each using an MCMC length of 500 million generations, sampling every 5000 iterations. Replicate runs were evaluated in TRACER v1.6 (Rambaut et al. 2013), and we used similar log likelihood and parameter distributions as indicators of convergence on the posterior distribution. Sampled trees from both replicates were combined in TREEANNOTATOR v2.4.5 into a maximum clade credibility tree (MCC) using a posterior probability limit of 0.95 and a burn-in of 50%.

2.3.4 Testing hypotheses of lineage divergence

We tested specific hypotheses of evolutionary history for our study system to determine if a model of lineage divergence based on geography, or phenotype, or a combination of both best fit our data (Figure 2.2). We used the program Bayesian Phylogenetics and Phylogeography (BPP) v3.3 in concert with BFDIVIDER to determine the marginal likelihood for each model across 16 Gauss-Legendre weights (Yang and Rannala 2010; Rannala and Yang 2017).

BPP is computationally intensive and it was not possible to analyze the full outgroup data set for these analyses. Therefore, for each model we used the same reduced dataset of 500 randomly selected loci, with 140 bp each, taken from 27 individuals: six *D.*

marmoratus from Nantahala, six *D. quadramaculatus* from Nantahala, six *D. marmoratus* from Pisgah, six *D. quadramaculatus* from Pisgah, and three *D. carolinensis* (Figure 2.2). This data set had the same quality filtering and clustering parameters described in the methods above, but had no missing data or uncalled bases, and included an extra *D. carolinensis* individual to lessen biases that can arise when unequal numbers of individuals are used for terminal branches. After initial exploratory runs, we settled on prior distributions for ancestral population size [$\theta = G(2,2000)$] and root age [$\tau = G(14,100000)$] that produced stable posterior distributions with effective sample sizes for parameters that were greater than 200. Each model was assigned the same prior distributions, with finetune parameters adjusted manually to keep posterior acceptance proportions between 0.2 and 0.8 (Rannala and Yang 2003). Models were run with a burn-in of 50,000 and a sampling frequency of 10, for a total of 500,000 samples, with each model run twice to check for consistency between runs. Convergence for values of θ , τ , and log likelihood were confirmed visually in TRACER (Rambaut et al. 2013). Bayes Factors (BF) were calculated using beta values as specified in the BPP manual.

2.3.5 Gene flow among lineages

Reconstructed phylogenetic histories can be influenced by both historic and contemporary introgression among lineages. Therefore, to assess if this might be a factor contributing to the phylogenetic reconstruction of relationships among resolved lineages in our study, we tested for signatures of migration between lineages by first using TreeMix v1.13 (Pickrell and Pritchard 2012) to identify previous migration events between five major clades recovered from the Beast tree (see Results). We created a

maximum likelihood tree in TreeMix and modeled up to six migration events between the clades. Pickrell and Pritchard (2012) recommend TreeMix to be used in concert with f -statistics to assess whether potential migration events result from admixture or incomplete lineage sorting, which can produce similar genetic signatures. Thus, we used the program F4 v0.9 in conjunction with fastsimcoal2 v2.6 to calculate the f_4 -statistic, which measures whether patterns of allele frequencies among four populations are characteristic of introgression events or could be the result of incomplete lineage sorting alone (Reich et al. 2009; Excoffier et al. 2013; Martin et al. 2015). We compared the f_4 -statistic calculated across 1000 simulated data sets for four clades identified in TreeMix as potentially involved in introgression events.

2.3.6 Climate niche modeling

Based on the results of the analyses outlined above, all *D. quadramaculatus* and *D. marmoratus* individuals were divided into two allopatric lineages (“Pisgah” and “Nantahala”), with little admixture between them. To determine if climate may be associated with their divergence, we performed a climate niche analysis on the Pisgah and Nantahala lineages. Owing to the bias inherent in using limited data points for climate niche modeling (Fourcade et al. 2014), we supplemented the locality data from this study with localities from GBIF and two previous studies: (Jackson 2005; Wooten and Rissler 2011). GBIF data was taken from individuals identified as *D. quadramaculatus*, removing all data prior to 1960 to increase the likelihood that locality information was current, and curating the resulting data to remove spurious localities. Points were mapped in QGIS v2.14, and all points found to be outside of the range of *D.*

quadramaculatus were considered errant and excluded. Duplicate localities were removed from all data sets. This resulted in a combined data set of 218 localities for Nantahala and 108 for Pisgah.

Because not all localities were genotyped for assignment to a particular population, we set the boundary between Pisgah and Nantahala to where genetic distance between populations was most distinct. To determine this boundary, we first used the pairwise Euclidean distances between individuals genotyped for this study to create a Gabriel connectivity network using the R package *adegenet* (Gabriel and Sokal 1969; Arnaud 2003). We then used Monmonier's algorithm to identify points lying along the geographic boundary between the two populations (Monmonier 1973; Manni et al. 2004). We assigned all non-genotyped individuals above the boundary to Pisgah and below the boundary to Nantahala for subsequent analyses.

We created ecological niche models using MAXENT v3.4.1 together with the species distribution packages *dismo* and *ENMTools* in R (Warren et al. 2010; Hijmans et al. 2017). MAXENT applies a general-purpose machine learning method based on maximum entropy to determine the probability that an organism will be present at a given location based on a set of climate variables. Climate data was extracted from WorldClim v1.4 bioclimatic layers at a resolution of $\sim 1 \text{ km}^2$ (Hijmans et al. 2005). The bioclimatic layers represent 19 biologically meaningful variables, such as temperature, precipitation, and seasonality, intended to be useful in species distribution modeling (Hijmans et al. 2005). While correlation between some of the variables was high, these correlated pairs differed between Pisgah and Nantahala, making it difficult to reduce correlation without compromising model quality since the same set of variables must be used to build both

ENMs to conduct comparisons between models. Hence rather than remove correlated variables, we adjusted regularization values in MAXENT to penalize parameters that might otherwise lead to model overfit (Phillips et al. 2006; Anderson and Gonzalez 2011; Elith et al. 2011)

When building ENMs, selection of the extent of background area to sample is a compromise between choosing an area large enough that the background points chosen do not overlap substantially with presence points, and small enough that the background encompasses an area that is ecologically relevant to the species, sometimes interpreted as being close to the boundaries of their geographical range (Barve et al. 2011). There is no consensus on how to select the optimal extent of the background area; thus, to minimize subjective bias we tested a range of background sizes (Jarnevich et al. 2017). We first randomly selected 1000 background points buffered to be a maximum of either 5, 10, 15, 20, 25, 30, 35, or 40 km from presence points (VanDerWal et al. 2009). We used the R package *ENMeval* (Muscarella et al. 2014) to evaluate all MAXENT models with k-fold cross-validation, partitioning data into four bins and assessing all combinations of models containing up to three feature classes (linear, quadratic, and product), and regularization values from 0.5 to 10 at intervals of 0.5 (Phillips and Dudík 2008; Merow et al. 2013). The best model was chosen as the one with the lowest AICc across all background sizes; note that AICc tends to favor models that are slightly underfit but are more robust to sampling bias (Galante et al. 2017).

We used the R package *ENMTools* to test hypotheses about whether the best fit ENMs for Pisgah and Nantahala were qualitatively different from each other, and hence likely to represent differences in niche use between the two lineages (Warren et al. 2010).

First, we conducted a symmetrical background test (also called a similarity test) to determine whether the overlap in habitat use for both lineages was due primarily to overlap in the available habitat (i.e., individuals in both lineages select the same habitat variables because the same variables are available in both areas; (Warren et al. 2008). We then used an identity test to compare whether the empirical values measuring niche overlap for the two lineages significantly differed from the expected distribution, which would indicate that occurrence of individuals in the landscape is not random (Warren et al. 2014). Finally, we ran a blob range-break test to determine whether the climates of the geographic areas occupied by Pisgah and Nantahala were significantly different from each other (Glor and Warren 2011). For each test, we compared two empirical measures of niche similarity, Schoener's D and Hellinger distance (I), against null distributions for each measure built from 1000 randomized replicate ENMs (Warren et al. 2008).

2.4 Results

We generated a total of 394,730,226 paired-end reads across 59 sequenced individuals. Average read depth per locus was ~56x across all loci and individuals. After data filtering, the ingroup data set had 3,305 total loci with 10% missing data (range = 2064–3284 loci/individual; $\bar{x} = 3029 \pm 271$), for a total of 476,453 bp and 21,764 parsimony informative sites. The outgroup data set had 3,206 loci with 12% missing data (range = 1398–3181 loci/individual; $\bar{x} = 2872 \pm 374$), for a total of 462,043 bp and 23,271 parsimony informative sites. Details of how locus filtering impacted the number of loci in the final data sets can be found in Table A2. All associated data (including links to raw sequence data) and input files for all analyses are available in Dryad.

2.4.1 Population structure

Cross validation of the ADMIXTURE results identified $K = 2$ as the optimal model of population structure, suggesting two distinct populations (Figs. 1C, A1). Individuals were not assigned according to their taxonomy, and instead were assigned with high probability (> 0.95) to clusters that correspond to geography. One cluster contained individuals sampled primarily from the southern portion of our study system in the Nantahala National Forest and the other cluster contained individuals sampled primarily from northern localities portion in the Pisgah National Forest. Hereafter, we refer to these units generally as Nantahala and Pisgah, respectively. A $K = 3$ model placed individuals from two localities (25, 26) in Georgia (GA), and three localities (11, 16, 17) in or near Great Smoky Mountains National Park (GSM), into a cluster separate from all other Nantahala individuals (Figure A2). Signatures of admixture between clusters were found in only six individuals at extremely low levels ($< 5\%$).

When individual *D. marmoratus* and *D. quadramaculatus* were assigned by species, DAPC analysis of the outgroup data used two PCs and 2 DAs, accounting for 71.6% of the conserved variance. All individuals from the two focal species were placed in the same region of discriminant space (Figure 2.3A), and distinct from *D. carolinensis*. Membership probability to either species varied between individuals, but only *D. carolinensis* individuals were correctly assigned with high probability to their putative species (Figure A3). Assignment of individuals to either Pisgah or Nantahala used eight PCs and 2 DAs, accounting for 84.7% of the conserved variance. The analysis resulted in two distinct clusters of individuals occupying opposite ends of the first discriminant axis (Figure 2.3B) with Nantahala and Pisgah individuals assigned with high probability to

their respective clusters (Figure A3). For comparison, *D. carolinensis* fell roughly in between the two populations in discriminant space.

F_{ST} between $K = 2$ clusters (Pisgah and Nantahala) was 0.76, while F_{ST} between both *D. carolinensis* and Pisgah, and *D. carolinensis* and Nantahala, was 0.90. For comparison, F_{ST} between *D. marmoratus* and *D. quadramaculatus* was 0.0004.

2.4.2 Phylogenetics

SVDQUARTETS analysis of the ingroup data produced a tree that resolved similar clustering patterns to those produced in ADMIXTURE (Figure A4). Individuals from Nantahala and Pisgah were placed in separate clades in the unrooted tree with 100% bootstrap support, and GA and GSM individuals were each placed in separate clades divergent from other Nantahala individuals. Analysis of the outgroup data set failed to place *D. carolinensis* in the tree in a position with greater than 50% bootstrap support and we do not present these results here.

Concatenated BEAST analyses of the outgroup data set resolved a deep divergence between a clade containing all Nantahala individuals and a clade containing Pisgah individuals and the *D. carolinensis* outgroup (Figure 2.3C). Within this latter clade, all Pisgah individuals were placed in a clade sister to *D. carolinensis*. All branches had posterior probabilities > 0.95 , except for the split between Pisgah and *D. carolinensis*, which was 0.82. Branch heights for deeper divergences in the Nantahala clade had large ranges, leaving the relative timing of these splits uncertain.

2.4.3 Testing hypotheses of lineage divergence

While there was not decisive support for a single model, our BPP results indicated that there is decisive support for the overarching hypothesis that geography is the main indicator of lineage divergence, not morphological phenotype (Table 2.1). Specifically, models where lineages were split primarily by geography (Figure 2.2, models B, C & D) had the highest support, while models where lineages were split primarily by phenotype (E & F) had Bayes Factors 1000x higher ($BF = 7418$ and 2748 , respectively). Results were consistent across replicates, with log marginal likelihoods varying by ≤ 5 units between replicates for any given model. The best-fitting model changed slightly between runs, with models B and C virtually indistinguishable in run 1, and models C and D tied in run 2. Placing *D. carolinensis* sister to the ingroup (models A & C), or sister to the Pisgah lineage (models B & D), had minimal to no effect on the marginal likelihood.

2.4.4 Gene flow among lineages

TREEMIX found 3 migration edges: Pisgah to Nantahala ($W = 0.032$, $P = 0.040$), *D. carolinensis* to GSM ($W = 0.001$, $P = 0.499$), and Pisgah to *D. carolinensis* ($W = 0.921$, $P = 0.09$; Figure A5). The observed f_4 -statistic for Pisgah, Nantahala, GSM, and *D. carolinensis* was smaller than the simulated f_4 -statistic (-0.0028 vs 0.0001 , respectively). However, the difference was not large enough to rule out that the observed f_4 -statistic was due to incomplete lineage sorting, rather than introgression.

2.4.5 Climate niche modeling

The best models according to AICc had background sizes of 30 km for both Pisgah (Figure 2.4A) and Nantahala (Figure 2.4B). A full summary of model comparisons across background selection points can be found in Table A3, and the location of the background and presence points used in each of the best models is shown in Figure A6. The three variables with the highest permutation importance for the Nantahala model were precipitation of warmest quarter (WorldClim variable BIO18), isothermality (BIO3), and precipitation during the driest month (BIO14). Precipitation during the warmest quarter was also of highest importance in the Pisgah model, along with temperature seasonality (BIO4), and precipitation of the wettest month (BIO13; Table A4).

Raster overlap was low ($D = 0.28$), indicating low similarity between the ecological niche models for the two groups. The symmetric background test showed high correlation between the available habitat across Nantahala and Pisgah, but significantly lower niche overlap than would be expected based on available habitat ($P = 0.002$; Figure A7A,B). The identity test indicated that there was a significant difference in niche space use between the two areas ($P = 0.010$; Figure A7C,D). Finally, the blob range-break test detected a significant range break, indicating a significant difference between the climates found in Pisgah and Nantahala ($P = 0.019$; Figure A8).

2.5 Discussion

The most striking result of this study is the rejection of a two-species model involving the evolution of highly differentiated morphologies being driven by ecological speciation in sympatry. Previous studies assumed *D. quadramaculatus* and *D.*

marmoratus represent evolutionarily distinct lineages based on their morphological differences—with the fully aquatic *marmoratus* phenotype preferring the center of the stream and the semi-aquatic *quadramaculatus* preferring the edges (Martof 1962). Our results clearly show that the primary divergence in this group is based on geography, not morphology. Indeed, the idea that *D. quadramaculatus* and *D. marmoratus* arose through ecological speciation may turn out to be a just-so story, with the traits that make *D. marmoratus* seem more aquatic having no adaptive significance at all (Lennox 1991). However, absence of evidence is not evidence of absence (Altman and Bland 1995); given the constraints of working on taxa with large genomes, whether these phenotypes have a genetic basis is still open to question (Keinath et al. 2015; Nowoshilow et al. 2018). For the remainder of this discussion we will consider in detail the factors that may be shaping both genetic and phenotypic patterns seen in this system, and the wider implications for our findings.

2.5.1 Possible mechanisms underlying morphological differentiation

Individuals of both morphological phenotypes were spread evenly throughout the geographic range and across the phylogeny, clustering neither in physical nor genetic space. Our models testing lineage divergence reject the two species as currently defined as representing independently evolving lineages, which hereafter we will refer to as the *marmoratus* and *quadramaculatus* phenotypes to avoid confusion with the traditional species designations.

We hypothesize there are two probable explanations for the lack of genomic divergence accompanying the phenotypic differentiation seen in this group. First, there

may be fixed genetic differences between the phenotypes that were not evident in our analyses. Behavioral and morphological phenotypes can be driven by changes at a small number of loci under strong divergent selection (Excoffier et al. 2009; Wolf and Ellegren 2017). These adaptive fixed polymorphisms can persist despite continued gene flow (Fitzpatrick et al. 2008; De La Torre et al. 2014), and form small genomic islands of divergence that are difficult to detect without data covering a large portion of the genome (Turner et al. 2005; Nosil et al. 2009). The loci sequenced here only covered ~0.005% of the ~15 Gb genome of these salamanders, making our data unsuitable for definitively detecting loci under selection (Storz 2005).

The second hypothesis is that the phenotypes are not genetically based, but the result of differential development. The two phenotypes might develop as a consequence of plasticity, possibly from environmental cues, which is a common phenomenon in amphibians [e.g. *Spea bombifrons* (Levis et al. 2018); *Ambystoma* spp (Jefferson et al. 2014); *Hynobius retardus* (Nishimura 2018)]. In *D. quadramaculatus*, plasticity during larval growth is correlated with abiotic variables, particularly temperature and rainfall (Camp et al. 2000). A higher density of *quadramaculatus* larvae are found in upstream tributaries compared to *marmoratus* (Bruce 1985), indicating that the two phenotypes may develop in environments that differ in water velocity, temperature, and/or predation risk (Rhoads and Sukhodolov 2001; Kiffney et al. 2006). If the phenotypes are a result of plasticity, habitats in regions where *quadramaculatus* is the sole expressed phenotype may lack environmental cues necessary to trigger development of the *marmoratus* phenotype. Environmental influences could either directly induce differential gene expression, leading to alternative phenotypes, or result in methylation modifications to

alter gene expression epigenetically (Jaenisch and Bird 2003). Epigenetically-mediated phenotypes can also be purely intrinsic, with genetically identical individuals producing heritable phenotypes (Wong et al. 2005).

Both phenotypes lay egg clutches on the undersides of rocks in streams and have a two- to three-year larval period (Martof 1962). However, the *marmoratus* phenotype is consistently smaller than the *quadramaculatus* phenotype at hatching and metamorphosis (Martof 1962; Bruce 1985). Differences in body size during development can have downstream effects on phenotype. For example, in the plethodontid genus *Thorius*, the shape of the eyes and nasal cavities are driven largely by mechanical interactions during morphogenesis, with smaller individuals subject to different mechanical pressures, thus developing different skull morphologies than their larger counterparts (Hanken 1983b,c). Consequently the features used to diagnose *D. marmoratus*—flattened head and snout, reduced internal nares, angled eyes (Dunn 1926)—could potentially be explained solely by cascading effects of body size on skull shape. Hence, environmental heterogeneity among and within streams could be driving different developmental pathways that explain the disconnect between phenotype and genetics seen in this system.

2.5.2 Geographic split corresponds to lineage divergence

Our hypothesis testing supported two geographic lineages—which we call Nantahala and Pisgah, after the national forests where they are largely located—with a strong history of divergence that we propose represents a previously unrecognized speciation event. High F_{ST} values and clear monophyly of both groups suggest that Nantahala and Pisgah have been isolated for an extended period of time. The paucity of

evidence for admixture between adjacent sites, sometimes even within the same stream, indicates that there may be a strong reproductive isolating mechanism preventing gene flow, creating a detectable genetic boundary between Nantahala and Pisgah.

A north/south divide within *D. quadramaculatus* and *D. marmoratus* has been suggested previously, with the French Broad River and the Eastern Continental Divide both proposed as likely candidates for a geographic boundary (Hinderstein 1971; Jones et al. 2006; Wooten and Rissler 2011). However, neither of these geological features align with the pattern of genetic divergence found here (Figure A9). Instead, the genetic boundary corresponds more closely with differences in climate—particularly precipitation levels, which are much higher in Nantahala than Pisgah (Figure A10). These climate differences do not follow any known geographic barriers that would prevent migration between the two areas, suggesting that climate may have played a role in divergence. Furthermore, the pattern we found may be an important component underlying the radiation of salamanders in the southern Appalachians, as it corresponds to a break point common to multiple species, including *D. carolinensis* (Tilley and Mahoney 1996; Kozak and Wiens 2006; Rissler and Smith 2010; Tilley et al. 2013).

One possible hypothesis is that the lineage divergence seen between Nantahala and Pisgah was driven by adaptation to different climate niches. Under this evolutionary scenario, the respective ranges of Nantahala and Pisgah would most likely have contracted during the climate extremes of the last 5-10 my as the availability of their ecological niches narrowed, leading to periods spent in allopatry without opportunities for gene flow to mitigate the effects of genetic drift. This narrative could explain the current lack of gene flow between adjacent individuals along the climate barrier.

However, the differences in the ecological niches of Nantahala and Pisgah may represent post-speciation evolution, with local adaptation to climate arising after reproductive isolation. We also cannot rule out the influence of other factors on divergence in this system (e.g. shifts in the position of drainage basins that could alter dispersal patterns) and admit that the dynamic topography of the area may lead to conflation of the effects of geography and climate, making it difficult to identify the specific agents involved in shaping genetic and ecological divergence in these lineages.

Given the strong genomic divergence between these two geographic lineages, it is particularly unusual that both lineages have maintained two divergent phenotypes. Phenotypic variation is common within many plethodontid species, with examples of color pattern polymorphisms (Fisher-Reid and Wiens 2015), variance in the number of tarsal and carpal bones (Hanken 1983a), and divergence of male secondary sexual characteristics (Sever 1979). However, we are unaware of an example of a plethodontid species with two morphotypes that each exhibit a set of distinct, consistently produced color pattern and skeletal traits. Furthermore, the presence of both morphotypes in two highly diverged lineages raises the question of whether phenotypic differences arose prior to the split between Pisgah and Nantahala and was then maintained through selective pressures (Pfennig et al. 2010; Arnegard et al. 2014; Lindtke et al. 2017), or if the phenotypes arose after the split via convergence (Renaut et al. 2014; Soria-Carrasco et al. 2014; Bailey et al. 2015). Both of these scenarios should produce a signature of genetic divergence that has thus far remained elusive, and acquisition of a data set with high enough coverage of the genome to detect signatures of selection will remain a challenge for this group until sequencing technology makes its next leap forward.

2.5.3 Broader phylogenetic perspective

The split of *D. quadramaculatus* into two geographically defined lineages is well-supported, but the position of these species within the broader desmognathine phylogeny remains tenuous. Our phylogenetic analyses in BEAST and SVDQUARTETS did not place Nantahala and Pisgah as sister clades, hinting at a more complicated evolutionary history with the rest of the genus. Our results potentially conflict with previous studies that resolved a *D. quadramaculatus* + *D. marmoratus* clade that was relatively divergent from *D. carolinensis* based on mitochondrial and a small number of nuclear markers (Kozak et al. 2005, 2009; Beamer and Lamb 2008). Our result may be influenced by our very limited taxon sampling across the genus, which could be affecting the placement of the root in our phylogenetic analyses, particularly when considering the relatively rapid radiation of the genus (Shavit et al. 2007). We do, however, note that our TREEMIX estimates of migration identified the greatest weight for gene flow as being from Pisgah to *D. carolinensis*, suggesting that introgression among lineages could be influencing our phylogenetic results (McVay et al. 2017; Morales and Carstens 2018). In any case, our results, coupled with the poor resolution among most of the major clades of *Desmognathus*, highlight the need for more comprehensive phylogenetic study across the genus with increased taxonomic and genomic sampling.

There are further questions posed by our findings that we were not able to fully address. For example, the deep division between GA, GSM, and Nantahala suggests that these areas may have experienced isolation in the past. Larger sample sizes from these regions, and better estimates on the timing of divergence, may aid in deciphering this issue. We also note there is a third species in this group, *D. folkertsi*, a dwarf form that

purportedly micropartitions the habitat while in sympatry with the Nantahala *quadramaculatus* phenotype (Camp et al. 2002, 2013b,a; Wooten et al. 2010). The fact that the range of *D. folkertsi* is completely encompassed within the range of Nantahala, and previous phylogenetic studies have produced contradictory results (Jackson 2005; Beamer and Lamb 2008; Kozak et al. 2009; Wooten et al. 2010), hints that *D. folkertsi* may have an equally complex underlying history as the rest of the group.

Our results join the growing number of examples where tests of putative ecological speciation revealed different—and more complex—histories hidden beneath the simple stories told by morphology alone. However, to resolve questions raised by our results in this system, further exploration of the role ecological differentiation plays in driving lineage divergence is needed. Specifically, modeling replicated patterns of divergence in co-distributed taxa across the climate differences we detected may provide insight into the broader scale divergence of the genus on both a temporal and spatial scale. While a comprehensive genomic study in this particular system may be challenging, we have shown here that coupling genetic and climate data can paint a more complete picture of evolutionary history, even in lineages traditionally recalcitrant to genomic methods.

Table 2.1. Results from replicate Bayesian Phylogeography & Phylogenetics. Runs for six models testing divergence by geography, phenotype, or a combination of both (see Figure 2.2 for more detail on models). Both phenotype models (E & F) had Bayes Factors much higher than any of the geographic models (A-D), indicating that the primary lineage split is based on geography, not phenotype.

| Model | Description | Log Marginal Likelihood | | Log Bayes Factor | |
|-------|-----------------------|-------------------------|---------|------------------|-------------|
| | | Run 1 | Run 2 | Run 1 | Run 2 |
| A | Geography only | -131841 | -131838 | 21 | 18 |
| B | Geography only | -131821 | -131826 | 1 | 6 |
| C | Geography + phenotype | -131824 | -131820 | 4 | best (tied) |
| D | Geography + phenotype | -131820 | -131820 | best | best (tied) |
| E | Phenotype only | -139238 | -139237 | 7418 | 7417 |
| F | Phenotype + geography | -134568 | -134565 | 2748 | 2745 |

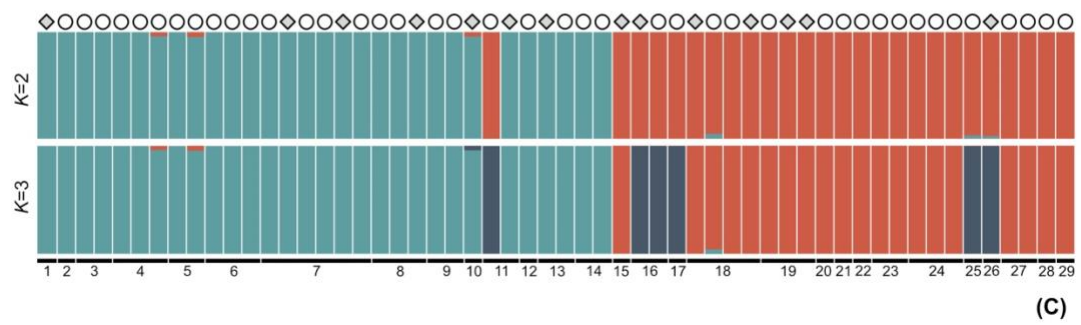
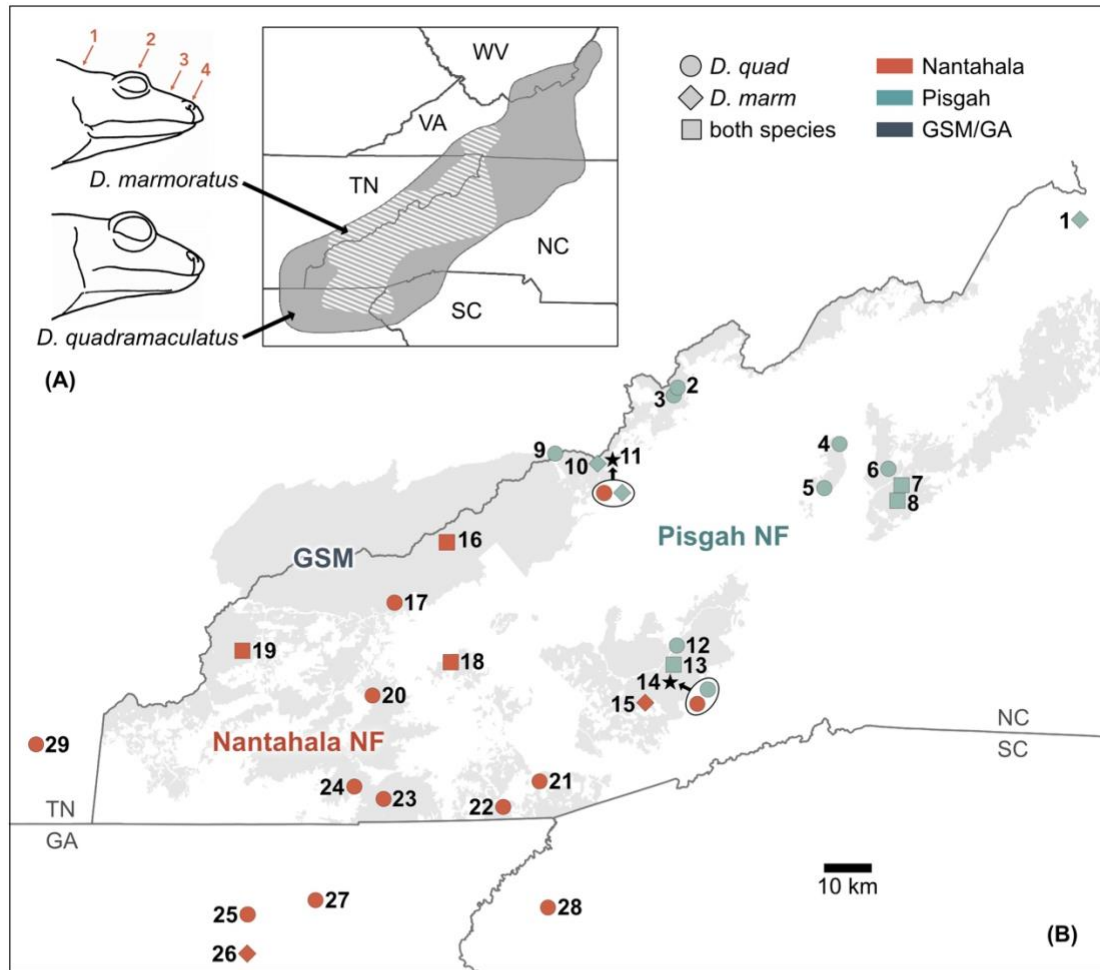


Figure 2.1. Sampling locations and phenotypic differences. (A) Differences in skull morphology for *D. marmoratus* (left): (1) steeper slope to the skull behind the eyes, (2) smaller eyes, (3) dorsoventrally compressed snout, and (4) different shape to internal and external nares. (Drawings adapted from Martof 1962.) Overlapping ranges for *D. quadramaculatus* (grey) and *D. marmoratus* (striped) across southern Appalachia (right). (B) Map of sampling sites for *D. quadramaculatus*, *D. marmoratus*, and both species within western North Carolina and adjacent states. Numbers on sampling sites correspond to Inset C and Table A1. Sampling sites are colored by Admixture results for $K = 2$. Forested areas colored light grey. (C) Results from Admixture for $K = 2$ arranged by sampling site, showing individuals grouping by geography rather than phenotype. Results for $K = 3$ can be found in Figure A2.

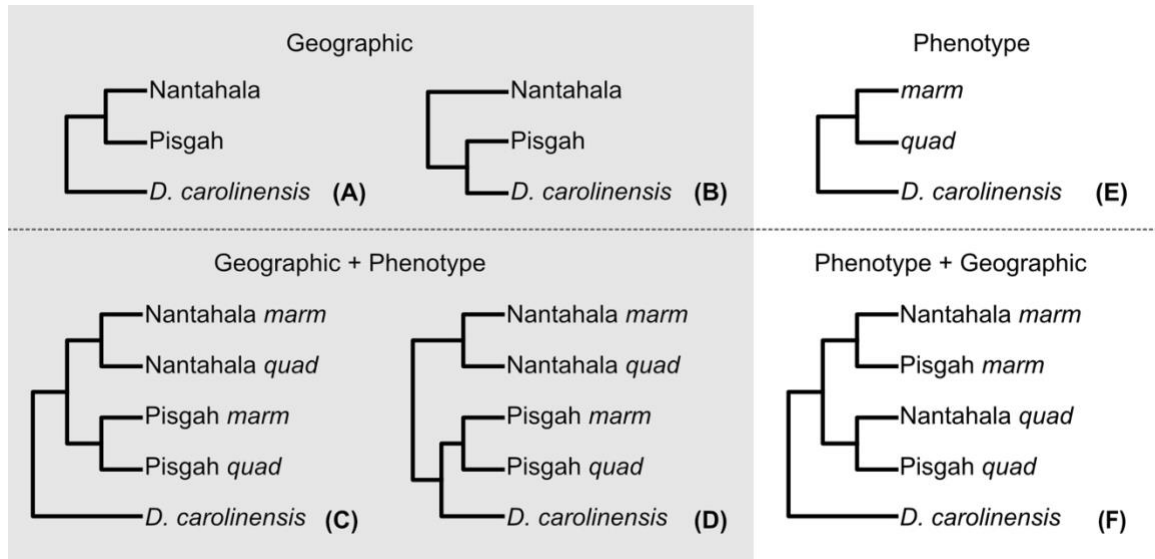


Figure 2.2. Models for Bayesian Phylogeography & Phylogenetics. Models for Bayesian Phylogeography & Phylogenetics to test whether the pattern of lineage divergence is based primarily on geography, phenotype, or a mixture of both. Geographic: genetic divergence follows geography only; (A) *D. carolinensis* is outgroup to *D. quadramaculatus* clade or (B) sister to Pisgah. Geographic | Phenotype: genetic divergence follows geography and then diverges by phenotype within geographic areas; (C) *D. carolinensis* is outgroup to *D. quadramaculatus* clade or (D) sister to Pisgah. Phenotype: (E) genetic divergence follows phenotype only. Phenotype | Geography: (F) genetic divergence follows phenotype and further diverged by geographic area.

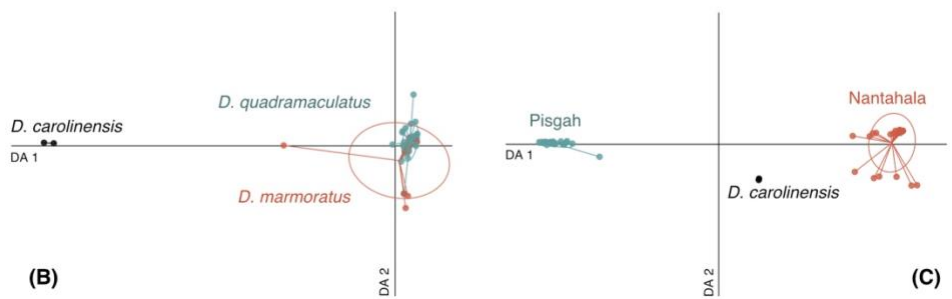
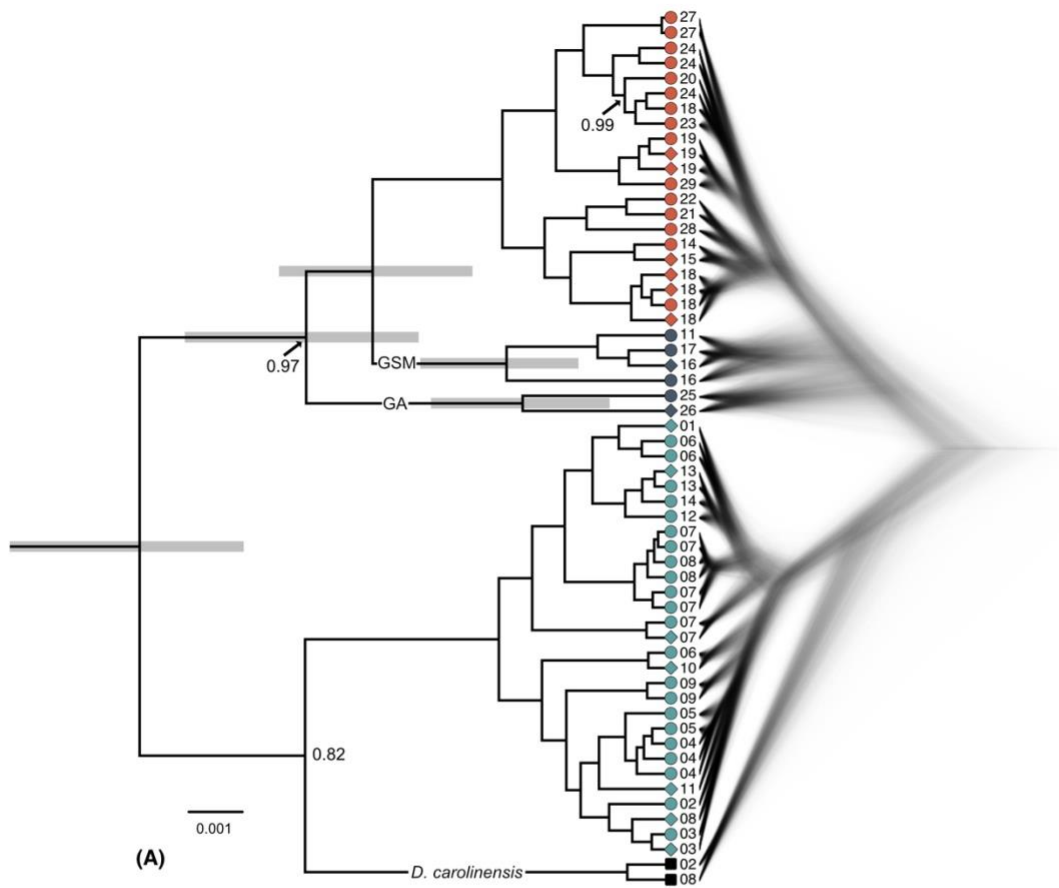


Figure 2.3. Results from Discriminant Analysis of Principle Components (DAPC). Individuals are shown as closed circles connected by lines to the centroid of their pre-assigned cluster membership, with each cluster surrounded by an ellipse representing the 95% confidence interval. (A) Individuals pre-assigned by phenotype, showing that the species as currently defined, *D. quadramaculatus* and *D. marmoratus*, could not be distinguished within discriminant space, but can both be distinguished from *D. carolinensis*. (B) Individuals pre-assigned by clusters identified in Admixture ($K = 2$), showing the geographic regions of Nantahala and Pisgah both forming distinct clusters on opposite sides of the second axis. (C) Results from Beast: maximum mean clade credibility consensus tree (> 0.95 posterior probability, 50% burn-in) for Pisgah, Nantahala, and *D. carolinensis* (left) and DensiTree using 50,001 trees (50% burn-in), with darker colors representing more common trees (right). The consensus tree covers 79.2% of all trees. Gray bars represent the range of tree heights with > 0.95 posterior probability. Shapes represent phenotypes for *D. quadramaculatus* (circles), *D. marmoratus* (diamonds), and *D. carolinensis* (squares). All nodes have a posterior probability of 1.0 except where noted. Colors represent Admixture assignments ($K = 2$). Note that Pisgah and Nantahala are paraphyletic, with *D. carolinensis* sister to Pisgah.

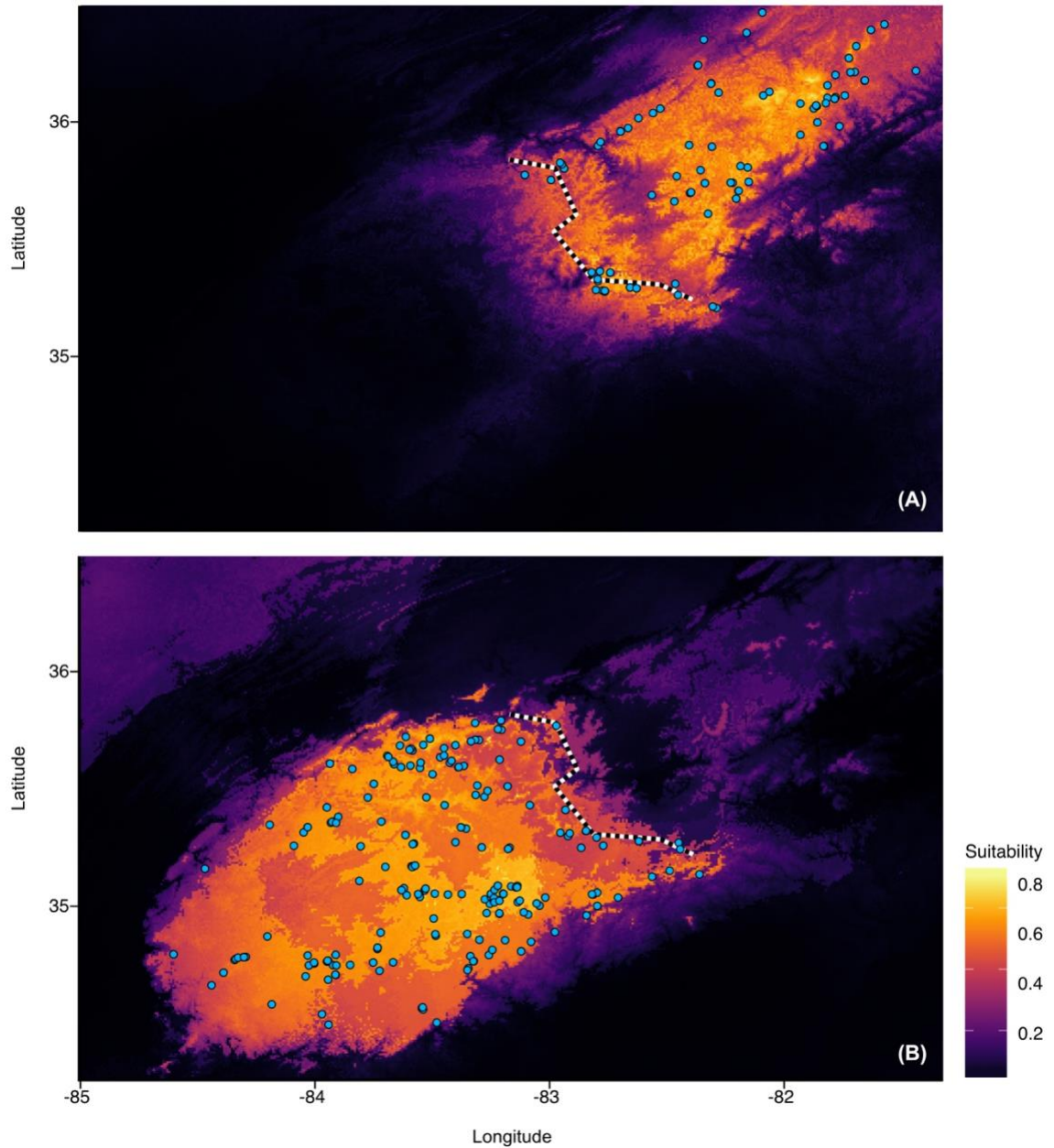


Figure 2.4. MaxEnt environmental niche models. MaxEnt environmental niche models showing habitat suitability across the same geographic area for (A) Pisgah and (B) Nantahala. Lighter colors indicate higher predicted suitability. Localities showing areas where *D. quadramaculatus/marmoratus* are present indicated with blue open circles. White grid lines show latitude and longitude. Black and white striped line indicates area of highest genetic difference between Pisgah and Nantahala and forms the putative boundary between these populations. There is very little niche overlap between the two groups.

CHAPTER 3. MITOCHONDRIAL AND NUCLEAR DATA SUPPORT LINEAGE DIVERGENCE WITHIN THE *DESMOGNATHUS QUADRAMACULATUS* SPECIES COMPLEX IN NANTAHALA BUT NOT PISGAH

3.1 Abstract

A number of studies have attempted to delimit species within the *Desmognathus quadramaculatus* species complex to better align taxonomy with the group's complex evolutionary history. However, there is disagreement between studies, which varied based on data type (i.e., mitochondrial, ddRAD, or anchored hybrid enrichment). Here, I attempt to reconcile these differences with a total evidence approach using publicly available data, in combination with a newly assembled draft transcriptome of *Desmognathus fuscus*. I used phylogenetic and population structure analyses, alongside haplowebs and conspecificity matrices, to determine if transcriptomic loci supported the hypothesis that the phenotypes represent multiple independently evolving lineages. I also used haplotype networks to re-examine whether phenotypes were associated with genetic divergence. No previously proposed candidate species had clear support in Pisgah, but there was support for genetic divergence between within Nantahala, though both clades had high levels of discordance due to a large number of uninformative loci. The weak population structure found in Pisgah is likely driven by isolation by distance, rather than admixture. These results reconfirm that morphology alone is not a good indicator of genetic divergence in this complex. Finally, the impact of systematic and stochastic error should be taken into consideration when interpreting phylogenetic and population genetic studies.

3.2 Introduction

In Chapter 2, I concluded that two members of the quadramaculatus complex of desmognathine salamanders, *Desmognathus quadramaculatus* and *D. marmoratus*, are not valid species as currently described and instead represent two divergent lineages that are geographically partitioned—Pisgah in the north of the range, and Nantahala in the south. I also found that variation in phenotype did not seem to be associated with genetic divergence. Since that chapter was published as Jones and Weisrock (2018), three studies were published that provide additional insight and raise new questions about the evolutionary history of this species complex. Beamer and Lamb (2020) performed a phylogenetic analysis using mitochondrial gene fragments (ND2 and COI) and proposed that seven clades within the complex warranted species-level status, with *D. folkertsi* (folkertsi) remaining as previously described, but *D. marmoratus* (marm) and *D. quadramaculatus* (quad) split into marmB, quadA, and quadF in Nantahala; and quad/marmC, quad/marmD, and quadE in Pisgah. Pyron et al. (2020) used anchored hybrid enrichment (AHE) to capture nuclear loci and expanded Beamer and Lamb (2020)’s mitochondrial data set to produce nuclear and mitochondrial phylogenies. They proposed that the four clades in Nantahala be considered four candidate species and that Pisgah be split into six candidate species, with the interdigitated marmoratus and quadramaculatus phenotypes elevated to separate candidate species. Finally, Pyron et al. (2022) expanded the AHE data set to include more individuals and used phylogenetic and population genetic analyses to propose a seventh candidate species in Pisgah.

While this wealth of data is potentially a boon for elucidating the evolutionary history of *Desmognathus*, closer examination of the analyses underpinning the

delimitation of eleven candidate species revealed several methodological issues which may have biased the results. For example, the PCA and SNMF analyses did not account for the impact of linkage disequilibrium or isolation by distance, which can create false clusters of individuals that resemble populations (Zou et al. 2010; Perez et al. 2018; Yi and Latch 2022). Most importantly, many of the clades proposed as candidate species had little to no branch support, and it is not known whether this low support was due to conflicting phylogenetic signals or a lack of signal in the data (Minh et al. 2020a). These issues suggest that careful re-analysis of the existing data might permit a more precise description of the evolutionary history of the group.

Thus, the purpose of this study was to determine whether the proposed candidate species are representative of the patterns of genetic divergence in the *Desmognathus quadramaculatus* species complex. To that end, I reassembled the AHE data from Pyron et al. (2022) for the Pisgah and Nantahala clades using a draft transcriptome of *D. fuscus* as a reference. I re-analyzed the resulting loci with the same methods used in that study with modifications to ensure adherence to model assumptions, and performed multiple replicates to evaluate the robustness of the results. I also added new analyses, such as haplowebs and conspecificity matrices, to gain insight into the complexities of the genetic data. Finally, I generated haplotype networks to explore mitonuclear discordance and phenotype switching in the focal clades. My goal was not to provide specific recommendations for taxonomic changes but to thoroughly interrogate the results to identify unresolved questions about the evolutionary history of these taxa.

3.3 Methods

3.3.1 Nuclear data preparation

The loci used in Pyron et al. (2022) (hereafter, Pyron22) were generated using anchored hybrid enrichment (AHE), which relies on a quasi-*de novo* assembly approach (Lemmon et al. 2012). To improve the accuracy of locus assembly, I reprocessed raw reads from Pyron22 using a draft transcriptome of *Desmognathus fuscus*—the most closely related species with an available transcriptome or genome—as a reference, as described below.

3.3.1.1 Raw sequence processing

I downloaded 150 bp paired-end reads in FASTQ format for *Desmognathus quadramaculatus*, *D. marmoratus*, and *D. folkertsii* from NCBI using the SRA Toolkit v3.0.0 (Leinonen et al. 2011). Individuals that were sequenced using single-end reads were excluded from the analysis (Table B1). Raw reads were trimmed and filtered using Fastp v0.23.2 (Chen et al. 2018). Adapter sequences were removed using autodetection (-detect_adapter_for_pe) and duplicate reads were removed using the highest accuracy setting (--dedup --dup_calc_accuracy 6). Bases were changed to Ns (i.e., missing) if they had a Phred quality score ≤ 30 or mean read quality dropped below 30 within a 4-bp sliding window from 3' forward. Reads were excluded if they contained five or more Ns, >20% of bases failed quality control, read complexity dropped below 30%, or read length was <100 bp after trimming.

3.3.1.2 Read mapping, variant calling, and candidate locus assembly

Trimmed reads were assembled with BWA-MEM using the *Desmognathus fuscus* transcriptome as a reference and output in BAM format (Li 2013). To control for paralogy, the transcriptome was filtered to include only the primary isoform for each gene, and reads were discarded if they mapped to more than one transcript or either mate in a read pair mapped to a different transcript. Alleles were called using BCFtools *mpileup*, with a maximum depth of 1000 and an indel bias of 0.75 (Danecek et al. 2021). All sites supported by fewer than ten informative reads (i.e., allele depth ≥ 10) were removed. Single nucleotide polymorphisms (SNPs) were removed using BCFtools *filter* if they were within 3 bp of another variant (-g3) and indels were removed if they were separated by 5 or fewer bp (-G5). Indels were left-aligned against the reference (BCFtools *norm*). Variant sites with a genotype probability of < 0.99 were set to missing (BCFtools plug-in *+setGT*). Variant and non-variant alleles that passed quality control were output into a single VCF (BCFtools *view*). A subset of VCF calls was visualized in the Integrative Genomics Viewer (IGV) alongside the BAM files used for allele calling to check that variant calls accurately reflected the composition of mapped reads (Robinson et al. 2017).

I built consensus loci with BCFtools *consensus* by applying alleles called in the VCF files to the reference transcriptome sequence (Danecek et al. 2021). Sites absent from the VCF (i.e., uncalled sites and sites that failed quality control) were marked as Ns. Leading and trailing missing bases were removed using a custom script. Only loci that had a minimum shared locus length of 400 bp across $\geq 50\%$ of individuals after alignment were retained as candidate loci. For loci that spanned multiple regions across one

transcript, only the longest locus was retained so that each locus putatively represented a single contiguous sequence. Scripts used for read mapping, variant calling, and locus assembly are available from the author upon request.

3.3.1.3 Final locus selection and data set construction

Due to the genetic distance between the Pisgah and Nantahala clades, some loci captured well across all individuals, while others had sparser coverage. Thus to minimize spurious assemblies and maximize the number of loci available for within-clade analyses, I created three data sets: 1) desmo900-all, containing only loci that were present in all individuals in both the Nantahala and Pisgah clades; 2) desmo900-pisgah, containing loci present in $\geq 50\%$ of individuals in the Pisgah clade; and 2) desmo900-nant, with $\geq 50\%$ of individuals in the Nantahala clade. For each data set, all loci were required to be a minimum length of 400 bp and contain at least one parsimony informative site. Candidate loci were aligned using MAFFT L-INS-i with a maximum of 1000 iterations (Katoh et al. 2002).

3.3.1.4 Comparison between desmo900-all and Pyron22 loci.

To determine how the loci compared between studies, the anchored hybrid enrichment loci used in Pyron22 were downloaded as individual PHYLIP files from the Dryad accession and converted to FASTA with Phylip2Fasta. For each locus, a single sequence from the Fuscus A clade was aligned to the *Desmognathus fuscus* transcriptome using SequenceServer 2.0.0 using *blastn* and an e-value cut-off of 1×10^{-20} (Priyam et al. 2019).

3.3.2 Mitochondrial data preparation

I also included mitochondrial data from two publicly available mitochondrial gene fragments, downloaded from NCBI. The first was a 2014 bp mitochondrial gene fragment consisting of the NADH dehydrogenase subunit 2, tRNA-Trp, tRNA-Ala, tRNA-Asn, tRNA-Cys, tRNA-Tyr, and cytochrome oxidase subunit I (COI) genes (Beamer and Lamb 2020). The second was a 527 bp fragment of the cytochrome b (CYTB) gene (Jackson 2005). Data was subset to include only the focal species, leaving 55 individuals in the COI data set and 376 individuals in the CYTB data set. After subsetting, both data sets were re-aligned using MAFFT as described above.

3.3.3 Genetic diversity, isolation by distance, and population structure

I used D_{PS} (i.e., allelic distance, or $1 - \text{the proportion of shared alleles}$) calculated between each pair of individuals to assess genetic variance within and between *a priori* groups (i.e., candidate species from Pyron22). D_{PS} provides a more sensitive metric for detecting reduced gene flow than more commonly used population statistics, such as F_{ST} (Landguth et al. 2010). I also calculated the number of diagnostic alleles, private alleles, and singletons using the *private_alleles* function in the R package *poppr*. For this study, diagnostic alleles were defined as fixed homozygous alleles found only in a single group (Kalinowski 2010); private alleles were exclusive to one group, but either not found in all members of the group or not fixed in all members of the group (Slatkin 1985); and singletons were minor alleles found in only one member of a group. The presence of isolation by distance (IBD) can lead to the overestimation of population structure due to the effects of genetic drift in organisms with limited dispersal (Aguillon et al. 2017; Perez

et al. 2018). Therefore I tested for IBD in Pisgah and Nantahala separately using a Mantel test (*mantel.randtest* in the R package *ade4*) with Euclidean genetic and geographic distance matrices over 999 permutations. To account for the tendency of Mantel tests to overestimate IBD when there is hierarchical population structure, I also tested *a priori* groups separately (Meirmans 2012).

3.3.4 Population structure

Most analyses used to detect population structure can only account for missing data by replacing it with zero values or the mean value, either of which can bias results (Yi and Latch 2022). Population structure can also be skewed by linkage disequilibrium and sites are expected to be independent (Zou et al. 2010; Frichot et al. 2014). Singletons are uninformative about group membership and can form specious clusters if they are present in high numbers in a data set (Linck and Battey 2019). I removed sites with missing data, thinned SNPs to one per locus, and removed sites with singletons from the data sets used in population structure analyses. To ensure that the results were robust and compare the effects of SNP selection on analyses, each analysis was run with eight replicates with one SNP per locus chosen randomly with replacement and the results between replicates were compared for consistency.

To determine whether individuals in each of the nuclear data sets clustered into groups, I performed principal component analyses (PCA) with the *dudi.pca* function of the *ade4* package in R, using normalized and centered variables (Jombart 2008). I also used sNMF as implemented in the *LEA* R package (version 4.2.0) to explore population structure in a model-based framework. sNMF uses sparse nonnegative matrix

factorization with least-squares optimization to estimate individual ancestry coefficients, producing results comparable to the program Admixture (Frichot et al. 2014). This program was also used in Pyron22, thus allowing a direct comparison of results between studies. I assessed values of K from 1-10, running 100 repetitions for each replicate SNP data set. Ancestry coefficients with the lowest cross-entropy score were kept for each value of K and each replicate. I used CLUMPAK to choose optimal cluster assignments for each replicate individually and to create a combined estimate of ancestry coefficients across all replicates (Kopelman et al. 2015). Results were visualized using the *pophelper* package in R (Francis 2017).

3.3.5 Phylogenetic analyses

I used IQ-TREE v2.2.0 to infer a partitioned concatenated maximum likelihood phylogeny and individual gene trees (Chernomor et al. 2016; Kalyaanamoorthy et al. 2017; Minh et al. 2020c). I assessed branch support with gene concordance factors (gCF), site concordance factors (sCF), and 1000 ultra-fast bootstraps (Hoang et al. 2018; Minh et al. 2020a). I used genesortR to further explore the effects of phylogenetic usefulness on branch support for the desmo900-all data set. This R script uses a PCA approach to concurrently assess multiple variables associated with systematic bias and phylogenetic signal including root-to-tip variance, level of saturation, average patristic distance, proportion of variable sites, average bootstrap support, and Robinson-Foulds similarity (Mongiardino Koch 2021). I sorted loci by decreasing phylogenetic usefulness, discarded the bottom 25% of loci, and used the remaining loci to infer a new phylogeny using the same parameters as above.

3.3.6 Haplotype networks, haplowebs, and conspecificity matrices

Unlike phylogenies, which assume that individuals bifurcate from an unobserved common ancestor, haplotype networks allow for direct connections (i.e., edges that represent one or more mutation events) between individuals with shared alleles (Carlo et al. 2001). Haplotype webs (a.k.a., haplowebs) are an extension of haplotype networks with additional curves added between haplotypes that co-occur in heterozygous individuals (Spöri and Flot 2020). The haploweb framework has been proposed as a method of species delimitation that groups individuals into “fields for recombination” (FFRs) that are considered distinct species because they share a mutually exclusive set of alleles and thus are considered to be reproductively isolated (Flot et al. 2010; Dellicour and Flot 2015).

For the mitochondrial data, I built haplotype networks by determining haplotypes for each individual using the *haplotype* function in R (*pegas* package) and connecting them into a TCS network in PopArt v1.7 (Paradis 2010; Leigh and Bryant 2015). TCS networks allow for multiple connections between haplotypes, which in the case of haploid markers can be interpreted as alternative connections as determined through statistical parsimony (Paradis 2018). For each of the desmo900-all loci, I used HaplowebMaker to determine haplotypes based on shared alleles and connected them using a median-joining network algorithm with an epsilon value of zero (Bandelt et al. 1999; Spöri and Flot 2020). Indels were treated as 5th characters (i.e., distinct alleles) and ambiguous characters were masked across all individuals. Finally, I used the partition matrix of FFRs generated by HaplowebMaker to calculate two conspecificity matrices for desmo900-all with the program CoMa: 1) using the sum of loci that support

conspecificity and 2) using the sum of loci supporting conspecificity minus the number of loci that support heterospecificity (Spöri and Flot 2020). All loci were given equal weight in the calculations. I clustered the resulting matrices with the *heatmap3* package in R, using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) agglomeration method (Zhao et al. 2014).

3.4 Results

3.4.1 Read mapping, variant calling, and locus assembly

For the desmo900 data set an average of 78% of reads mapped to the transcriptome, with 75% primary mapping, and 35% mapping in proper pairs. After filtering, there were 221 assembled loci with a mean alignment length of 1590 ± 722 bp (min length = 424 bp, max = 4311 bp) across 159 individuals in the desmo900-all data set. For the desmo900-nant set, 235 loci were recovered across 53 individuals, with a mean alignment length of 1539 ± 746 bp (424-4311 bp). For the desmo900-pisgah set, 258 loci were recovered across 106 individuals, with a mean alignment length of 1475 ± 739 bp (414-4311 bp). Further statistics for the data sets are included in Table 3.1.

Out of the 233 loci used in Pyron22, 210 (90.1%) matched at least one sequence in the transcriptome, with 196 (93.3%) of those loci having a >95% identity match with a transcript. Of the Pyron22 loci that matched a transcript, 124 (59.0%) matched more than one transcript and 93 (44.3%) had transcripts that overlapped with one other. Out of the 235 loci in the desmo900-all data set, 171 (72.7%) were built from transcripts that matched portions of Pyron22 loci. Thus while there is a large overlap in the data used in

Pyron22 and this study, not all of the loci are directly comparable because some of the Pyron22 loci were composed of fragments from more than one transcript.

3.4.2 Genetic diversity, isolation by distance, and population structure

Genetic distance varied widely between groups defined in Pyron22, with the highest genetic distance found between individuals in Nantahala being almost twice as large as that found in Pisgah (8.6% vs 4.5%, respectively; Table 3.2). The highest between-group genetic distance was between a pair of individuals in quadD and quadE (3.9%) in Pisgah and quadA and quadF (7%) in Nantahala. The highest within-group genetic distance did not exceed the lowest between-group diversity for any group except when comparing all of Pisgah to all of Nantahala.

There were 703 diagnostic alleles over 180 loci separating Nantahala from Pisgah (Table 3.3). All groupings tested had a high number of private alleles (range: 493-7830), with 28-58% of private alleles being singletons. All groups within Nantahala had diagnostic alleles (range = 44-189) regardless of whether individuals were grouped into two or four clades. Only two groups in Pisgah had diagnostic alleles (marmE = 6 and quadD = 2).

The Mantel test showed no significant IBD in Nantahala ($R = 0.15$, $p = 0.003$). Pisgah as a whole had significant IBD ($R = 0.54$, $p = 0.001$), and all the clades in Pisgah had significant IBD except for marmE ($R = 0.45$, $p = 0.079$). Regression plots of genetic distance versus geographic distance showed that individuals within Nantahala partitioned into discrete clusters, while individuals within Pisgah formed a continuous cline (Figure 3.2). However, a local mean smoothing curve fitted against the data showed that the

relationship between genetic and geographic distance in Pisgah was non-linear, indicating that IBD likely contributes to genetic structure but does not fully explain it.

The PCA for desmo900-all clustered individuals from Pisgah and Nantahala on either side of the first axis, but could not consistently differentiate individuals from within Pisgah or Nantahala from each other. When assessed separately, Nantahala split into three clear clusters (quadA, quadF, and folkertsi+marmB) in all replicates. For Pisgah, 5/8 replicates split quadD+quadE from all other individuals along the first axis, while all other groupings were inconsistent between replicates (Figure 3.3).

For the Nantahala sNMF analyses, $K = 4$ recovered the four clades found in Nantahala, with little shared ancestry between them. $K = 3$ was similar, but folkertsi and marmB had a high amount of shared ancestry. At $K = 5$ and above folkertsi shared ~25% of ancestry with marmB, and quadA split into two clusters with various amounts of shared ancestry between individuals, indicating that $K = 5$ likely overfits the data (Figure 3.4).

For Pisgah, $K = 2$ clustered individuals at the opposite ends of the geographic range, with shared ancestry for quadG and marmE (which are more centrally located). $K = 3$ split marmE into a cluster, with quadG sharing ~50% of ancestry with marmE and marmC, and quadC sharing some ancestry with marmE as well. $K = 4$ split quadD into a cluster, though some individuals in quadE shared >25% ancestry with quadD, and marmC shared more ancestry with marmE (Figure 3.5).

3.4.3 Phylogenetic analyses

The genesortR analysis found a phylogenetic usefulness axis that explained 24.54% of the variance between loci. When sorted by phylogenetic usefulness, the

bottom 25% of loci tended to have the highest root-to-tip variance, highest site saturation, lowest proportion of variable sites, lowest bootstrap support, and lowest Robinson-Foulds similarity (Figure 3.6). The topology was consistent between the phylogeny built with all loci and with only phylogenetically informative (PI) loci. Exclusion of the bottom 25% of loci resulted in slightly higher gCF and sCF support for most branches, indicating that the excluded loci were likely adding more noise than signal when used for phylogenetic inference.

For both phylogenies, the branches with the highest support were those leading to the Nantahala clade (gCF = 95/97, sCF = 95.3/95.4, for all loci/PI only trees respectively) and Pisgah clade (gCF = 88.2/92.8, sCF = 96.5/97.1; Figure 3.7). Within Nantahala, there were moderate gCFs and high sCFs for all four clades: folkertsi (gCF = 49.3/53, sCF = 84.5/86.8), marmB (gCF = 62.4/69.3, sCF = 94.7/97.3), quadA (gCF = 61.1/69.9, sCF = 96.7/91.7), and quadF (gCF = 61.1/72.9, sCF = 85.5/97). Within Pisgah, all *a priori* groups had gCFs < 2, except for marmE (gCF = 7.7/6.63, 65.4/66.1). All branches within Pisgah had gDFp (gene discordance factor due to polyphyly) values > 90, indicating that the low gCF values were driven by noise rather than alternate gene tree topologies. However, Pisgah branches also had moderate sCF values (49.0-66.1), indicating that the decisive sites for each branch had relatively high discordance (Table 3.4). All branches described above—regardless of gCF or sCF—had bootstrap values of 100.

3.4.4 Haplotype networks, haplowebs, and conspecificity matrices

Both mitochondrial haplotype networks showed quadF connecting Nantahala to the northernmost portion of Pisgah. In the CYTB network, folkertsi and marmB had the

closest connection with quadA, while in the COI network those relationships were obscured by alternative edges (Figures 3.8 & 3.9). CYTB showed multiple instances of the marmoratus phenotype in Pisgah sharing a haplotype with quadramaculatus, but each phenotype in the COI network had its own set of haplotypes (though one marmB individual was nested within quadA).

Out of 221 haplowebs, 12 (5%) were uninformative for distinguishing between Pisgah and Nantahala (Figure 3.10). For groups within Pisgah, 182 (82%) had no haplotypes that were exclusive to any of the groups (i.e., all groups shared at least one haplotype with an individual from another group), 11 had haplotypes exclusive to marmE, 7 had haplotypes exclusive to quadD, and one had a haplotype exclusive to quadG. An additional 5 loci had haplotypes that were exclusive to quadD, but the quadD haplotypes were not clustered together. For Nantahala, 117 (53%) showed each of the four groups having exclusive haplotypes, 26 had shared haplotypes between marmB and folkertsi, 9 shared between quadA and folkertsi, 2 shared between quadF and folkertsi, 1 shared between quadF and marmB, and 13 shared haplotypes between more than two groups.

The two conspecificity matrices differed in their support for candidate species. The matrix built using only the sum of loci supporting conspecificity showed support for delimiting Pisgah as a single species (Figure 3.11). Folkertsi, marmB, quadA, and quadF each had high conspecificity, but there was also (weaker) support for folkertsi and marmB being grouped together and there was some overlap between individuals in quadA and quadF. For the matrix built using the sum of loci supporting conspecificity minus the sum supporting heterospecificity, the results were far less clear. The overall

groupings remained similar due to high rates of heterospecificity between groups, but there was low support for all the individuals within any group being delimited together as one species due to conflicting signals between genes (Figure 3.12).

3.5 Discussion

3.5.1 Genetic structure in Pisgah is likely driven by drift

It is difficult to assess with certainty the degree to which locus construction, adherence to model parameters, and SNP selection contributed to the differences in sNMF results because Pyron22 only reported a single value of K for Pisgah ($K = 7$) and Nantahala ($K = 5$), and the values of K used in Pyron22 produced results that are difficult to interpret (Lawson et al. 2018). For example, Pisgah is split into seven clusters in Pyron22, but one of those clusters is only represented by a single individual (not included in this study) that shares ancestry almost equally with the six other groups regardless of geographic distance, which is not biologically plausible. There are also varying amounts of shared ancestry between the groups in both the Pisgah and Nantahala analyses which may be the result of model overfitting from using a higher value of K than supported by the data (Novembre 2016). Furthermore, even at higher values of K , this study did not recover many of the same patterns of admixture used to support candidate species in Pyron22, such as a splitting quadE into two north/south populations (labeled E1 and E2 in Pyron22).

Population structure plots can be interpreted hierarchically, with the lowest values of K representing the strongest signal (Lawson et al. 2018). In the sNMF plot for $K = 2$, the two most geographically distant groups (marmC in the south and quadD in the north)

are shown as two clusters, and individuals in geographically intermediate positions share ancestry with both clusters (Figure 3.4). This could be interpreted as admixture between the two most geographically distant groups, with the geographically intermediate individuals representing a ~200 km wide hybrid zone. However, when individuals along a genetic gradient (as expected with IBD) are forced into a pre-specified number of clusters, it will produce the same trend seen in the sNMF plots (Meirmans 2012).

The PCA replicates also show a pattern more consistent with IBD than admixture. Admixed individuals are expected to align in intermediate positions along the PCA axis separating their parent groups (McVean 2009). Instead, individuals appear to form a longitudinal gradient across the first and second axes, as expected for populations exhibiting IBD (Novembre and Stephens 2008). There were also no consistent clusters for marmC and quadD between PCA replicates, and only weak clustering for quadD+quadE from the rest of the groups in Pisgah (Figure 3.3E & F). Thus the PCA results indicate that drift due to IBD is more likely the driver of population structure than admixture and that high values of K in the sNMF analysis do not represent separate populations.

The paucity of diagnostic alleles found in the Pisgah groups provides further evidence that admixture is not the reason that some individuals in Pisgah show shared ancestry between groups. Diagnostic alleles are genetic synapomorphies that arise when mutations become fixed in a population due to selection or genetic drift. Their absence in most Pisgah groups indicates they have not been genetically isolated long enough for fixation or that there are a large number of F1 hybrid individuals in the analyses. F1 hybrids would be heterozygous at potentially diagnostic sites—which would mean their

inclusion in the analysis would “hide” potentially diagnostic alleles—but after the first generation, backcrossed hybrid individuals should have recoverable diagnostic alleles from their parent populations (Kalinowski et al. 2011; Malde et al. 2017; Coster et al. 2018). However, even under the relaxed assumption that 20% of marmC are F1 hybrids with another group, there would still only be two diagnostic alleles for the group, compared to 50 for folkertsi. If the sNMF results reflect admixture proportions between groups, rather than IBD, then diagnostic alleles should be recoverable from each group (della Croce et al. 2016).

In addition to IBD, there is support for other processes affecting genetic structure in Pisgah. All levels of clustering in the sNMF analyses show genetic breaks for individuals around the Asheville Basin (Figure 3.4). This is most apparent in the Blue Ridge Mountains southwest of Asheville, and the Black Mountains, where there is less shared ancestry between individuals that are geographically close but on the opposite side of the mountain range (<5 km apart) than with geographically distant individuals on the other side of Asheville Basin (>40 km apart). There are several possible explanations for this phenomenon. One is that the genetic break is an overinterpretation of the sNMF results (i.e., $K = 1$ is a better representation of the data than $K = 2$) and therefore spurious (Lawson et al. 2018). However, a similar pattern was also seen in the Black Mountains for $K = 3$ in Chapter 2, indicating that the high peaks in the Black Mountains may be acting (or acted in the past) as a physical or ecological barrier to gene flow. Similarly, gene flow north and south of the Asheville Basin may be constrained by the lack of suitable habitat within this low-lying area, leading to higher genetic distances between

individuals when the dispersal routes rejoin southwest of the Basin, following the pattern of an “ephemeral” ring species (Bouzid et al. 2022).

The results also cannot rule out that past demographic events or isolation may have contributed to genetic structure, particularly for quadD and marmE, both of which had multiple diagnostic SNPs and exclusive haplotypes. Furthermore, marmE is the only group in Pisgah that does not show significant IBD, indicating that it has a unique evolutionary history. The pattern of quad+quadE clustering away from the rest of Pisgah in the PCAs is also intriguing because it could indicate that gene flow between the central and northern portions of Pisgah was reduced in the past. However, there are too few independent markers to effectively explore these trends with the currently available data.

3.5.2 Some of the candidate species do not form well-supported clades

Despite the differences in locus construction, the topology of the concatenated trees produced in this study were comparable to those in Pyron22, except that in this study quadC and quadG do not form reciprocally monophyletic clades but instead form a clade with marmC. Though the authors in Pyron22 do not explicitly define which species concept they are using, they state that “we treat reciprocal monophyly of geographically distinct clades as the clearest evidence for valid candidate species.” In the case of Pisgah specifically, they claim that the “distinctiveness of the candidate species is supported by their formation of genealogically exclusive clades in the concatenated phylogeny.” Hence, their definition of a candidate species relies on whether a group is well-supported as a monophyletic clade, and the validity of the candidate species hinges on whether there is support for the clades.

Branch support was determined using three metrics: bootstrap support, gene concordance factors (gCFs), and site concordance factors (sCFs). All the branches leading to clades that were defined as candidate species had bootstrap values of 100%. Bootstrap values are calculated by resampling from the underlying data to determine the proportion of pseudosamples that support a branch (Felsenstein 1985; Minh et al. 2013). However, in large data sets sampling variance decreases while systematic error does not, resulting in high bootstrap support even for branches that are not well supported by the underlying data (Kumar et al. 2012; Thomson and Brown 2022).

In contrast, concordance factors provide detail on how many genes or sites support a branch. It is not unusual for gene trees to lack concordance (Lozano-Fernandez 2022), and processes such as incomplete lineage sorting can lead to discordant signals between gene trees, especially for rapidly evolving species with short branch lengths (Wiens et al. 2008). Furthermore, the topology that most accurately reflects the evolutionary history of a group may not be the most common topology found among gene trees (Degnan and Rosenberg 2006; Wang and Hahn 2018). Therefore, it is important to consider not just whether a branch is in concordance with a majority of genes or sites, but what underlying processes are contributing to discordance and what these processes can tell us about the evolutionary history of the branch.

Gene concordance factors can range from 0-100%, with 0% indicating that there is either a high degree of discordance between genes or that no genes support the branch, and 100% indicating that all genes support a branch. However, gCF is further broken down into three statistics: gDF1, gDF2, and gDFp (Minh et al. 2020b). The first two statistics describe the proportion of genes that support two alternative branch topologies

(i.e., they quantify discordance between gene trees), while the final statistic describes the proportion of genes that do not support any of the tested topologies (i.e., gene discordance due to polyphyly). For Nantahala, gCFs ranged from 49.3 to 65.2, with most of the discordance arising due to gDFp, meaning that about half of the gene trees provided support for the clades being monophyletic, while the other half were not informative about the branch. For Pisgah, the highest gCF was only 7.7, with over 90% of the discordance due to gDFp, indicating that the lack of branch support was due to a lack of a single clear signal from any of the gene trees, rather than conflicting signals among different gene trees.

Site concordance factors resample sites across all loci and test site support using randomly generated quartets from the resampled data. Since quartets can produce three possible topological arrangements around a branch, a randomly generated alignment is expected to produce an sCF of ~33%, and an sCF of 50% indicates that half of the quartets tested support alternate arrangements. The sCFs for Nantahala were high (84.5-96.7%) and thus support the Nantahala clades as being monophyletic. The sCFs for Pisgah clades ranged from 47.1 to 66.1 in this study. However, sCFs represent a proportion of the number of decisive sites, which varies between branches. For example, the branch with the highest support in Nantahala (quadA) had ~496 decisive sites, ~480 (96.7%) of which supported the clade. In Pisgah, the clade with the highest support was marmE, which had ~212 decisive sites, ~140 sites (65.4%) of which supported the clade, and ~73 that supported alternative topological arrangements. Hence, the clades in Pisgah had fewer sites that supported each branch, and those sites did not produce concordant results.

Overall, the phylogenies indicate that the Nantahala clades have support for being reciprocally monophyletic, and potentially valid candidate species under the definition used in Pyron22. However, there is not enough information in the data to determine with certainty whether the candidate species in Pisgah are reciprocally monophyletic due to a lack of clear phylogenetic signal. This is unsurprising considering that the branches leading to groups in Pisgah are approximately half the length of those supporting groups in Nantahala, and short branches are difficult to resolve (Kapli et al. 2020). If branch lengths represent an accurate estimate of divergence time, it also suggests that Pisgah is younger than Nantahala (Lanfear et al. 2010; Ho 2014). This interpretation is also consistent with the smaller genetic distances and higher number of shared haplotypes seen in Pisgah.

3.5.3 Haplotype networks are consistent with concordant mitonuclear divergence between Pisgah and Nantahala

Previous phylogenies inferred from ND2 and COI mitochondrial genes show Nantahala and Pisgah as sister clades that form a monophyletic group (Kozak et al. 2005; Beamer and Lamb 2020; Pyron et al. 2020), whereas in the AHE phylogenies the two clades are separated by *D. aeneus* and *D. imitator* (Pyron et al. 2020, 2022). Pyron et al. (2020) proposed this mitonuclear discordance arose when an ancient hybridization event transferred mitochondria from one clade to another. This is not unprecedented within *Desmognathus*; a mitochondrial introgression event between *D. carolinensis* and *D. fuscus* is well-documented (Kratovil 2017). However, there are several aspects of the data that warrant additional consideration when assessing the introgression hypothesis over alternative explanations regarding mitonuclear discordance.

First, it is important to consider the accuracy of the placement of *D. aeneus* and *D. imitator* in the *Desmognathus* phylogeny. Both of these clades have long branches, which can bias tree topology through a phenomenon known as long branch attraction (LBA), whereby groups with high numbers of substitutions cluster together away from their “correct” places on the tree (Lozano-Fernandez 2022). The placement of *D. aeneus* and *D. imitator* varies depending on the mitochondrial loci used in phylogenetic inference (Rissler and Taylor 2003; Jackson 2005; Kozak et al. 2005; Beamer and Lamb 2020) and large nuclear data sets tend to exacerbate LBA rather than resolve it (Susko and Roger 2021). It is therefore possible that the two clades share a most recent common ancestor and the insertion of species between Pisgah and Nantahala in Pyron22 is an artifact of LBA.

Second, if Pisgah and Nantahala evolved independently followed by introgression, then we would expect to see discordance between the mitochondrial and nuclear trees within the clade that received the introgressed mitochondria. This is the pattern seen with *D. carolinensis* and *D. fuscus*—individuals descended from that introgression event group with *D. carolinensis* in mitochondrial phylogenies and *D. fuscus* in nuclear phylogenies. However, mitochondrial and nuclear trees that were inferred using the same individuals are concordant for the same patterns of genetic divergence within Pisgah and Nantahala (Pyron et al. 2020), suggesting that mitochondrial and nuclear divergence occurred concurrently.

Third, both mitochondrial haplotype networks show a single connection between Pisgah and Nantahala via quadF (in the Nantahala clade). Surprisingly, the haplotypes in the COI network that showed the most similarity were between those in quadF and

quadD, which is the most geographically distant group in Pisgah from quadF. The pattern is less clear for CYTB, but still shows a closer relationship between individuals near the Great Smoky Mountains in Nantahala and above the Asheville Basin in Pisgah, rather than with those it overlaps with geographically.

Together, these observations suggest that the introgression hypothesis may not be the best fit for the available data and suggest an alternative hypothesis: Pisgah diverged from Nantahala when northern portions of the population became isolated, followed by recolonization when ecological conditions improved (e.g., after the Last Glacial Maximum). This hypothesis would be concordant with (1) similar patterns of genetic divergence within mitochondrial and nuclear analyses; (2) the pattern of divergence seen in the haplotype networks and haplowebs; (3) the lack of hybridization between individuals where Pisgah and Nantahala overlap, despite their phenotypic similarity; and (4) lower genetic differentiation in Pisgah, despite being spread across a wider geographic range than Nantahala. Regardless of the specific scenario that led to divergence, Pisgah and Nantahala sharing a most recent common ancestor would also provide a more parsimonious explanation for phenotypic (e.g., quadramaculatus and marmoratus morphology) and ecological (e.g., long aquatic larval stage) similarities between the clades.

3.5.4 Revisiting the ecological role of phenotypes in divergence

In Chapter 2, I concluded that the quadramaculatus and marmoratus phenotypes probably did not play a role in genetic divergence. Pyron et al. (2020) asserted that the phenotypes form reciprocally monophyletic clades and that my previous findings were skewed by the inclusion of misidentified individuals. (Note that Chapter 2 only covered

two of the three nominal species and a smaller portion of the overall range than those studies.) However, some of the candidate species proposed in Pyron et al. (2020) and Pyron²² have little support in the analyses shown here, suggesting that the role of phenotype in divergence is still uncertain. Additionally, the inclusion of only marmoratus phenotypes in marmC may be a product of incomplete sampling; the CYTB phylogeny shows both marmoratus and quadramaculatus phenotypes in the southernmost reaches of the Pisgah clade (Jackson 2005). This is not to say that there is no relationship between phenotype and divergence in Pisgah, only that the relationship is based on conflicting evidence and requires further research to resolve.

It can be difficult to assess the role of ecology in phenotypic divergence when a species covers a large geographic area because phenotypic differences spread along an environmental gradient could be a result of ecological adaptation or drift under IBD (Seeholzer and Brumfield 2018). Genetic drift during bottlenecks can also result in fixation onto a single phenotype, mimicking the signatures of selection (Ferchaud and Hansen 2016; Ashraf and Lawson 2021). Hence genetic drift alone could result in the fixation of alleles responsible for the marmoratus phenotype. For example, if the portion of the Pisgah clade located west of the Asheville Basin is exclusively composed of one phenotype (marmoratus) and we assume that the Pisgah and Nantahala clades diverged in allopatry, then those westernmost individuals would represent the leading edge of the expansion of the Pisgah clade into sympatry with Nantahala. Hence, the switch from the more common quadramaculatus phenotype to the rarer marmoratus phenotype could be driven by selection for the smaller marmoratus phenotype when invading streams already occupied by the larger quadramaculatus phenotype (i.e., microhabitat partitioning), or it

could be the result of random fixation on the marmoratus phenotype due to founder effects.

Another important consideration is that the quadramaculatus and marmoratus phenotypes—along with the wide range of skin pattern variation associated with these morphologies (Pyron and Beamer 2022)—may result from polymorphism providing a selective advantage within a population. Skin patterns found across the quadramaculatus complex likely act as cryptic camouflage against the heterogeneous background of the second- and third-order streams where they spend most of their time. Pattern and body shape polymorphisms can reduce predation by disrupting the search image of predators (Murali et al. 2021; Troscianko et al. 2021) and allowing some individuals to remain cryptic despite spatial or temporal changes to the background (e.g., seasonal leaf fall) within suitable microhabitats (Polo-Cavia and Gomez-Mestre 2017; Hantak and Kuchta 2018; Burgon et al. 2020). Polymorphism may also allow populations to rapidly adapt to match local background colors and thus be advantageous to maintain across a species complex despite ongoing genetic divergence unrelated to phenotype (Jamie and Meier 2020; Barnett et al. 2021). Hence, rather than reflecting genetic divergence, phenotypic diversity could help maintain species boundaries by allowing one species to maintain the adaptive potential or phenotypic plasticity needed to thrive in a varied and changing environment. Further research to determine whether there are genotypic markers associated with phenotypes—and if they evolved independently—is necessary to fully understand the role that phenotypes play in the diversification of this group.

Table 3.1. Genotyping rate and data set sizes for nuclear loci used in this study.

| Data set | # ind. | # loci | Alignment length | Genotyping rate | | # SNPs | |
|--------------|--------|--------|------------------|-----------------|---------------|-----------------------|---------------------|
| | | | | All sites | Variant sites | Missing data excluded | Singletons excluded |
| desmo900-all | 159 | 221 | 351,424 | 99.6% | 99.3% | 17,320 | 11,317 |
| -nant | 53 | 235 | 361,794 | 99.2% | 97.4% | 11,901 | 7,960 |
| -pishah | 106 | 258 | 380,527 | 97.1% | 91.2% | 13,555 | 9,118 |
| PI only | 159 | 166 | 292,077 | — | — | — | — |

Table 3.2. Summary of pairwise genetic distance (D_{PS}) for individuals within groups and between groups. Between-group statistics calculates the lowest and highest pairwise distance between individuals in the group listed below it. Pisgah South includes individuals in groups marmC, quadC, quadG, and marmE; Pisgah North includes individuals from quadD and quadE. Note: groups in Pisgah marked with asterisks do not form reciprocally monophyletic clades.

| | Within-groups | | Between-groups | |
|-------------------------------|---------------|---------|----------------|---------|
| | lowest | highest | lowest | highest |
| Nantahala vs Pisgah | | | | |
| Nantahala | 0.2% | 8.6% | 12.0% | 15.2% |
| Pisgah | 0.1% | 4.5% | | |
| Nantahala major clades | | | | |
| folk+marmB | 0.2% | 5.5% | 6.9% | 8.6% |
| quadA+F | 0.8% | 7.0% | | |
| Nantahala minor clades | | | | |
| folkertsi | 0.2% | 2.0% | 3.5% | 5.5% |
| marmoratus B | 0.2% | 3.5% | | |
| quad A | 0.8% | 3.2% | 1.6% | 7.0% |
| quad F | 0.8% | 4.3% | | |
| Pisgah major clades | | | | |
| Pisgah South | 0.1% | 3.9% | 2.6% | 4.5% |
| Pisgah North | 0.5% | 3.9% | | |
| Pisgah minor groups | | | | |
| marmC | 0.1% | 1.7% | 1.6% | 2.6% |
| quadC* | 0.6% | 2.6% | | |
| quadG* | 1.3% | 2.6% | 2.6% | 3.5% |
| marmE | 0.4% | 2.6% | | |
| quadE | 0.4% | 3.2% | 2.9% | 3.9% |
| quadD | 0.7% | 3.0% | | |

Table 3.3. Number of private, singletons, and diagnostic alleles found within groups. Diagnostic alleles are defined as fixed homozygous alleles found exclusively in one group. Private alleles are exclusive to one group but not found in all individuals within the group and/or not fixed across individuals (i.e., heterozygous). Singletons are private alleles found only in a single individual. Pisgah South includes individuals in quadC, quadG, marmC, and marmE. Pisgah North includes quadD and quadE.

| | Nantahala | | Pisgah | |
|-------------------------------------|-----------|--|--------|--|
| Number of individuals | 53 | | 106 | |
| Diagnostic alleles | 703 | | 703 | |
| Number of loci w/diagnostic alleles | 180 | | 180 | |
| Private alleles | 7830 | | 7601 | |

| | folkerti+marmB | quadA+quadF | Pisgah South | | Pisgah North | |
|-------------------------------------|----------------|-------------|--------------|--|--------------|--|
| Number of individuals | 14 | 39 | 55 | | 51 | |
| Diagnostic alleles | 130 | 44 | 0 | | 0 | |
| Number of loci w/diagnostic alleles | 84 | 30 | 0 | | 0 | |
| Private alleles | 1513 | 4834 | 2668 | | 2961 | |

| | folkerti | marmB | quadA | quadF | marmC | quadC | quadG | marmE | quadD | quadE |
|--|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Number of individuals | 6 | 8 | 33 | 6 | 16 | 21 | 9 | 9 | 13 | 38 |
| Diagnostic alleles | 87 | 105 | 177 | 189 | 0 | 0 | 0 | 8 | 7 | 0 |
| Number of loci w/diagnostic alleles | 50 | 80 | 94 | 88 | 0 | 0 | 0 | 6 | 2 | 0 |
| Private alleles | 751 | 1210 | 4053 | 1644 | 493 | 1332 | 857 | 729 | 1752 | 2257 |
| Singletons | 350 | 763 | 2130 | 698 | 293 | 643 | 537 | 525 | 997 | 1442 |
| % of private alleles that are singletons | 53% | 37% | 47% | 58% | 41% | 52% | 37% | 28% | 43% | 36% |

Table 3.4. Gene concordance factors and site concordance factors. Comparison of gene concordance factors and site concordance factors for branches leading to monophyletic groups in the concatenated phylogeny for all loci assembled for this study, and the values previously published in Pyron et al. 2022. gCF = gene concordance factor; gDF1 = proportion of trees discordant factor for NNI-1 branch; gDF2 = proportion of trees discordant factor for NNI-2 branch; gDFp = proportion of trees discordant due to paraphyly; sCF = site concordance factor; sDF1 = site discordance factor for the first alternative quartet; sDF2 = site discordance factor for the second alternative quartet; sN = number of sites decisive for a branch. Bootstrap values were 100% for each branch shown. Note: quadG and quadC are not monophyletic and therefore gCF and sCF are not shown for those groups.

| | This study | | | | | | | | Pyron et al. 2022 | |
|-----------------|------------|------|------|------|------|------|------|----------|-------------------|------|
| | gCF | gDF1 | gDF2 | gDFp | sCF | sDF1 | sDF2 | sN | gCF | sCF |
| Nantahala | 95 | 0.5 | 0.5 | 4.1 | 95.3 | 2.4 | 2.3 | 1702.528 | 83.4 | 88.7 |
| folkertsi+marmB | 67 | 0.9 | 3.2 | 29 | 94.4 | 2.1 | 3.5 | 569.688 | 82.5 | 91 |
| folkertsi | 49.3 | 3.2 | 6.3 | 41.2 | 84.5 | 7.7 | 7.8 | 214.991 | 48.6 | 80 |
| marmB | 62.4 | 0.9 | 1.8 | 34.8 | 94.7 | 3.5 | 1.7 | 356.126 | 80.5 | 97 |
| quadA+quadF | 36.7 | 10.4 | 18.6 | 34.4 | 59.2 | 16 | 24.8 | 276.843 | 41.2 | 57.1 |
| quadA | 65.2 | 0.9 | 0 | 33.9 | 96.7 | 1.7 | 1.6 | 496.037 | 64.6 | 95.1 |
| quadF | 61.1 | 5 | 4.1 | 29.9 | 85.5 | 10.4 | 4.1 | 406.987 | 66.8 | 85.6 |
| Pisgah | 88.2 | 0 | 0 | 11.8 | 96.5 | 1.6 | 1.9 | 1351.247 | 78.5 | 96.6 |
| quadD+quadE | 0 | 0 | 0.5 | 99.6 | 53.3 | 15.5 | 31.1 | 243.024 | 1.3 | 49.2 |
| quadD | 0.9 | 0 | 0 | 99.1 | 66.1 | 20.7 | 13.2 | 193.776 | 2.2 | 54.9 |
| quadE | 0 | 0 | 0.9 | 99.1 | 47.1 | 19.5 | 33.4 | 168.931 | 0 | 50.5 |
| rest of Pisgah | 0.5 | 0.9 | 0 | 98.6 | 49 | 20.8 | 30.2 | 262.632 | 0.9 | 48.6 |
| marmE | 7.7 | 0 | 0 | 92.3 | 65.4 | 12.5 | 22 | 212.571 | 6.6 | 64.1 |
| marmC | 0.9 | 0.9 | 0.5 | 97.7 | 57.5 | 14.9 | 27.7 | 134.126 | 0.4 | 58.4 |

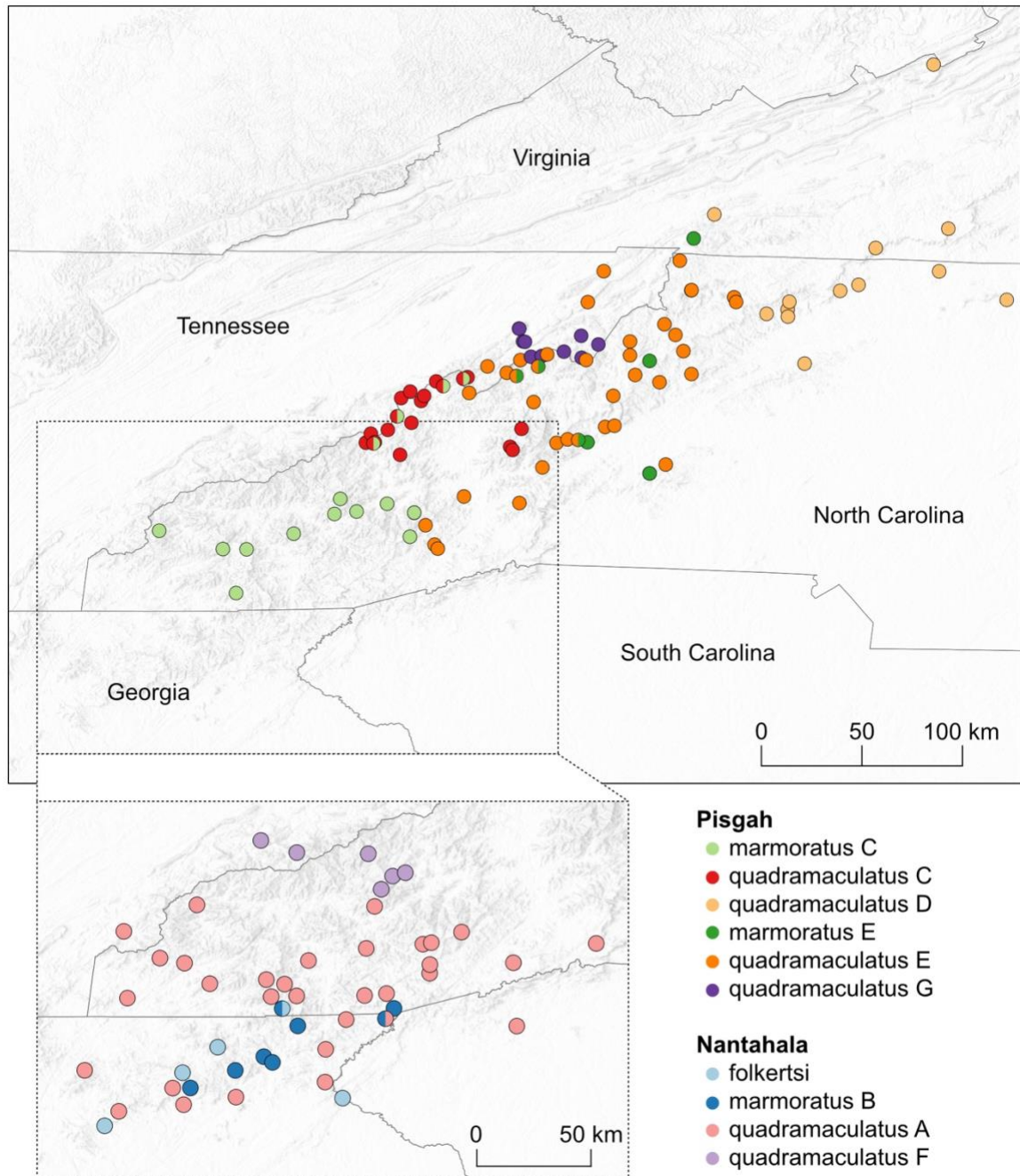


Figure 3.1. Sampling map of individuals used in this study. Pisgah clade (above) and Nantahala clade (inset, below). Dashed outline shows where Pisgah and Nantahala overlap. Individuals are colored after the candidate species in Pyron et al. (2022). Sites where individuals from two groups were collected are colored with half circles.

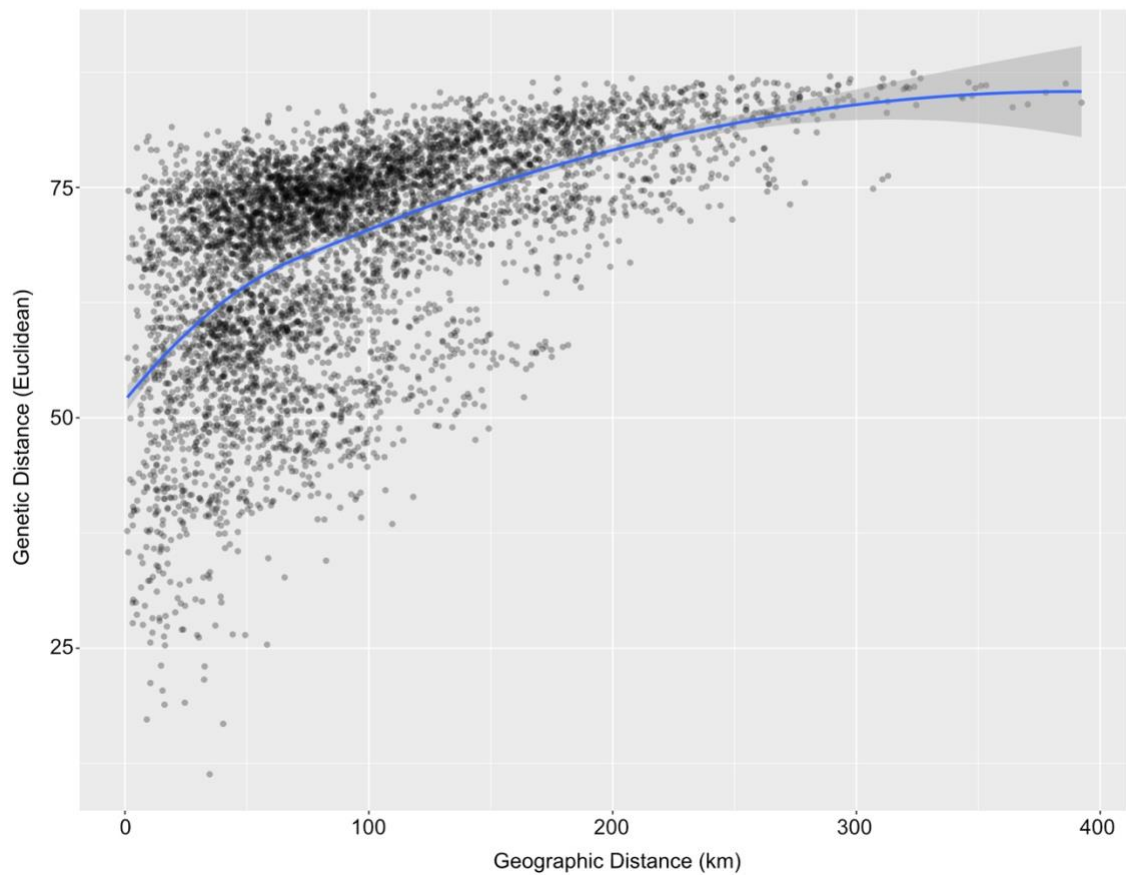
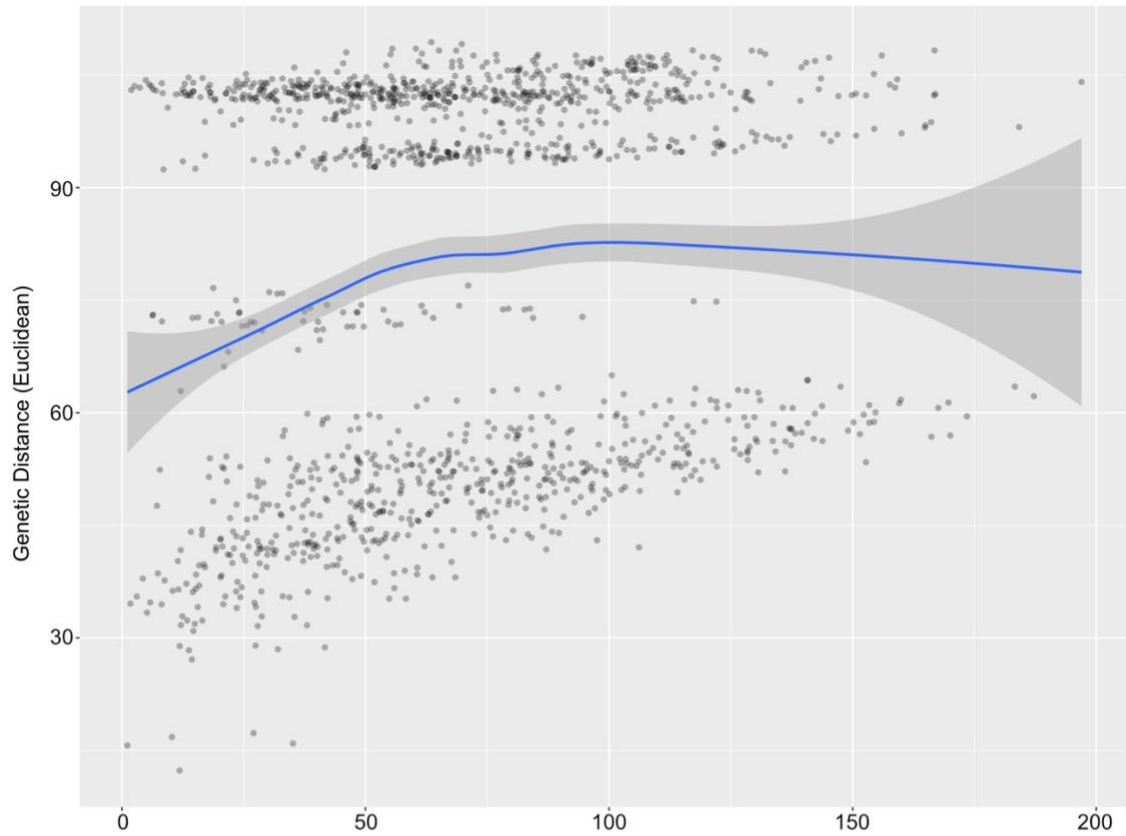


Figure 3.2. Euclidean genetic distance versus geographic distance for all individuals. Nantahala (top) and Pisgah (bottom), overplotted with smoothed local mean (blue curve) with a 95% confidence interval (grey shaded area).

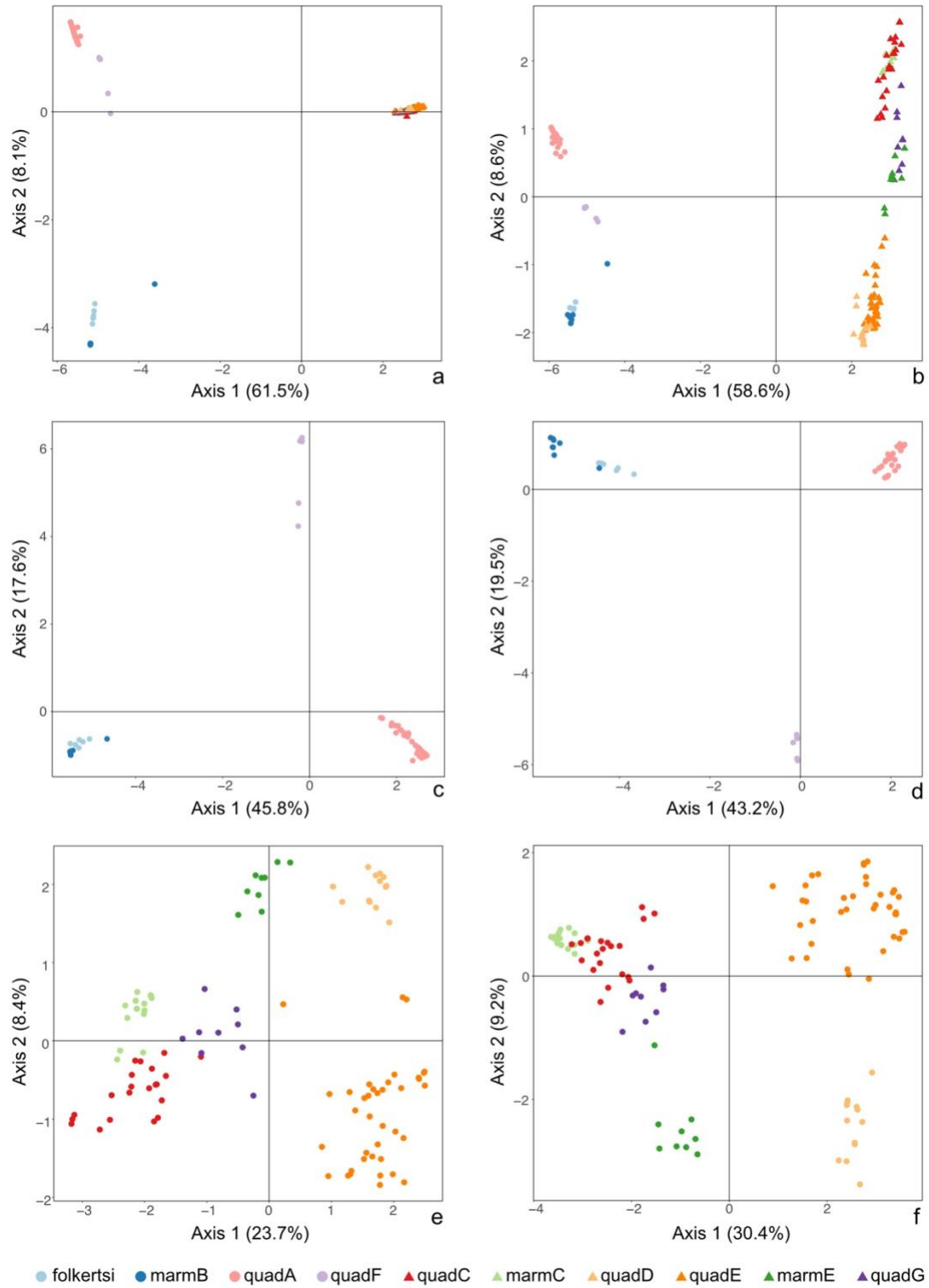


Figure 3.3. Principal component analysis. PCA for two replicates each of desmo900-all (a and b), -nant (c and d), -pishah (e and f). Individuals are colored based on groupings in Pyron22 for comparison. Percent variance explained by each axis included in parentheses.

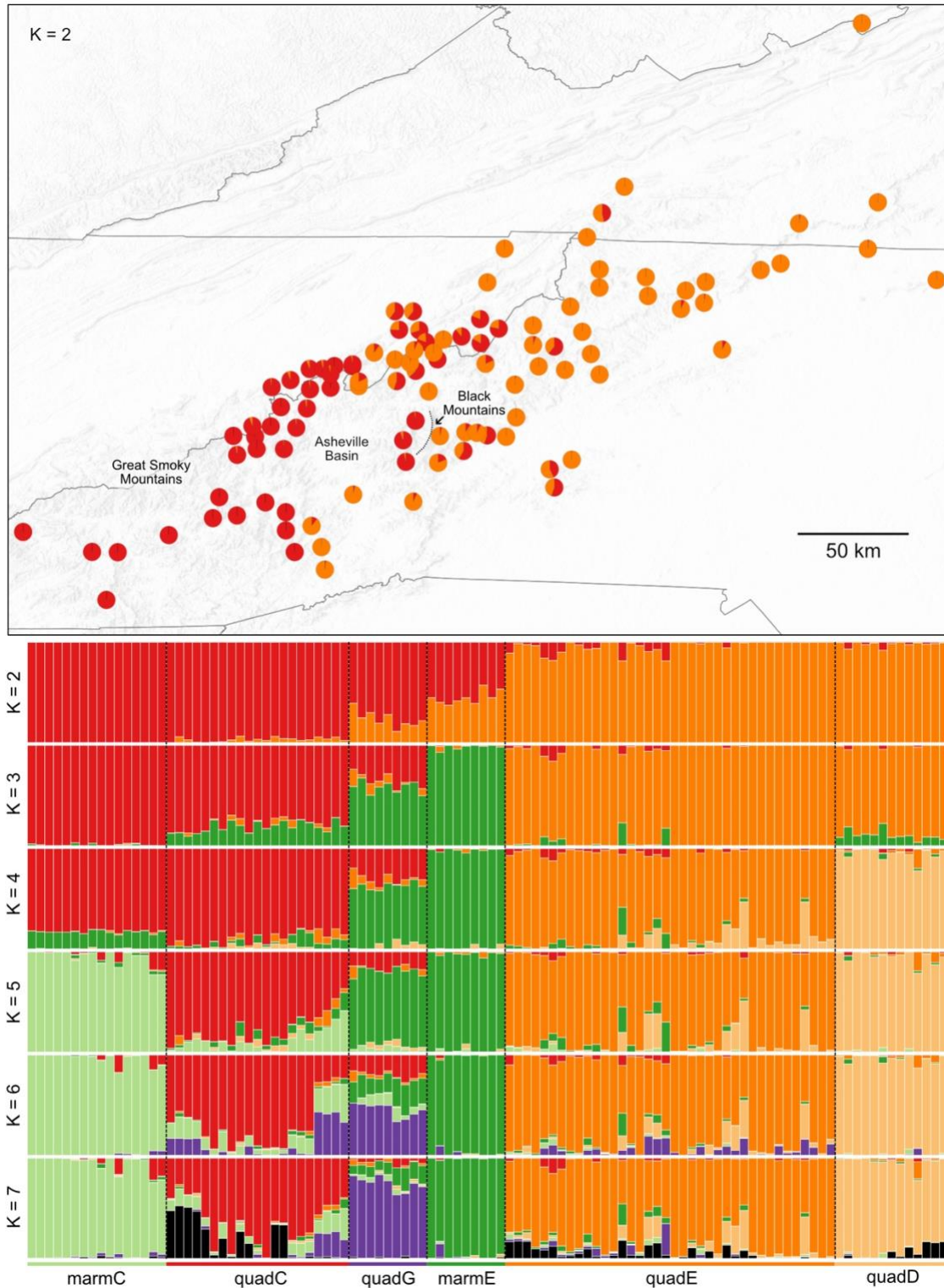


Figure 3.4. Population structure results from sNMF for Pisgah. Map (top) shows $K = 2$ and plots (below) show ancestry coefficients for $K = 2-5$. Individuals in sNMF plots are ordered roughly by longitude. Sites of interest mentioned in the text (Great Smoky Mountains, Asheville Basin, and Black Mountains) are indicated on the map.

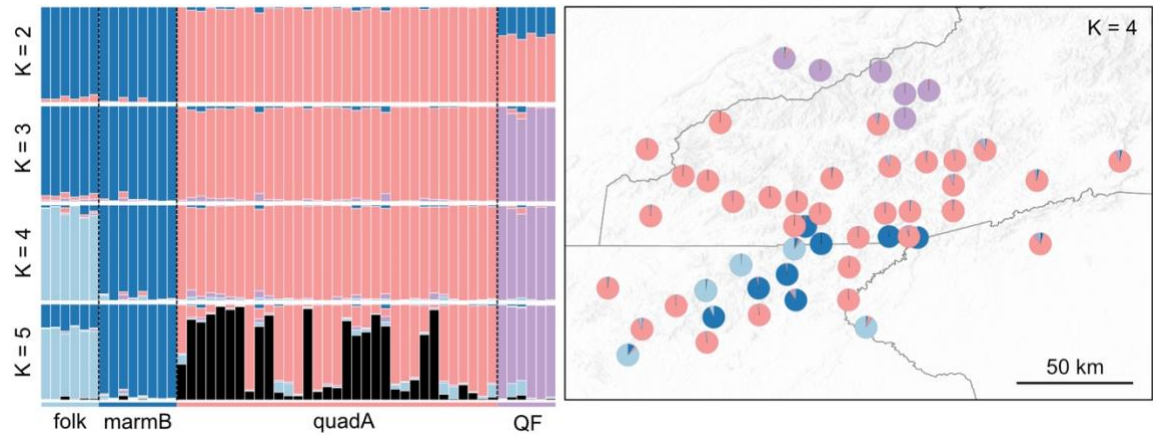


Figure 3.5. Population structure results from sNMF for Nantahala. Map (left) shows $K = 4$ and plots (below) show ancestry coefficients for $K = 2-5$.

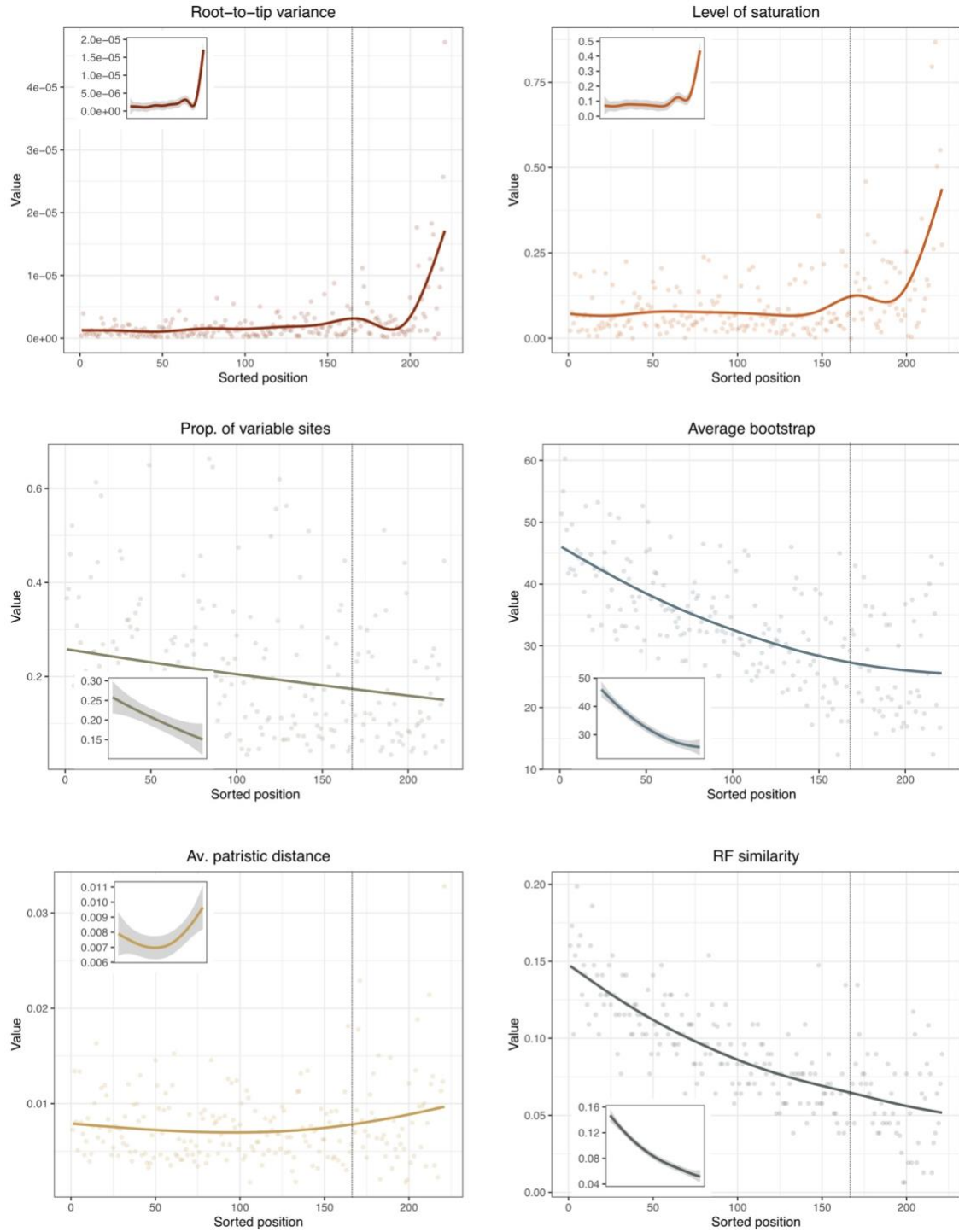


Figure 3.6. Six gene properties used to assess phylogenetic usefulness in genesortR for the 221 loci used in the desmo900-all phylogeny. Vertical dashed line indicates the cut-off for the bottom 25% of loci that were excluded from the tree inferred using the most phylogenetic informative loci.

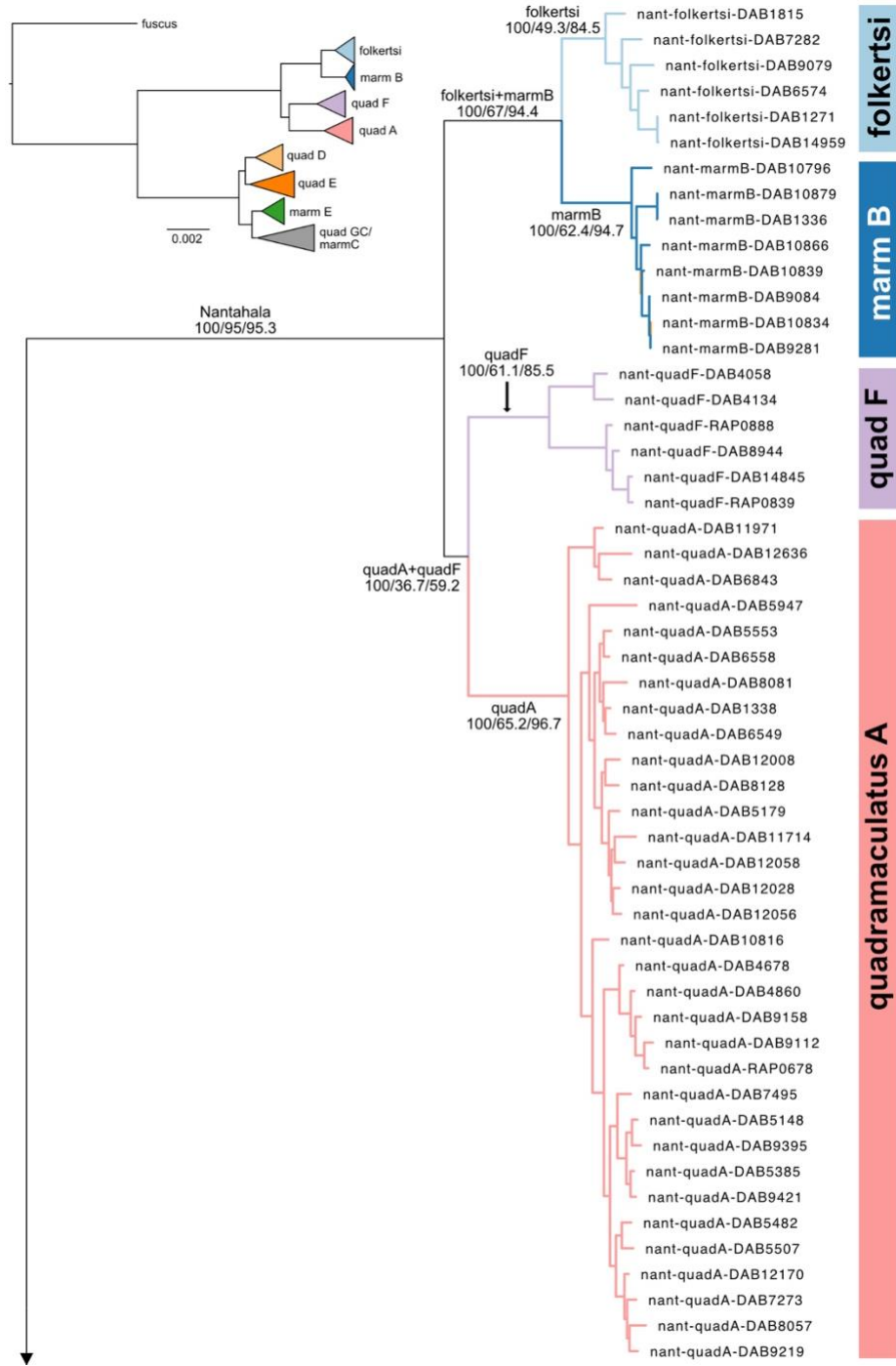


Figure 3.7. Maximum likelihood phylogenetic tree built using a partitioned concatenated matrix of desmo900-all loci. Bootstrap/gCF/sCF values shown for branches that represent clades and/or *a priori* groups from Pyron22. Branch colors and labels correspond to colors used in population structure analyses. Top left inset shows an overview of the same tree with individuals collapsed by clade.

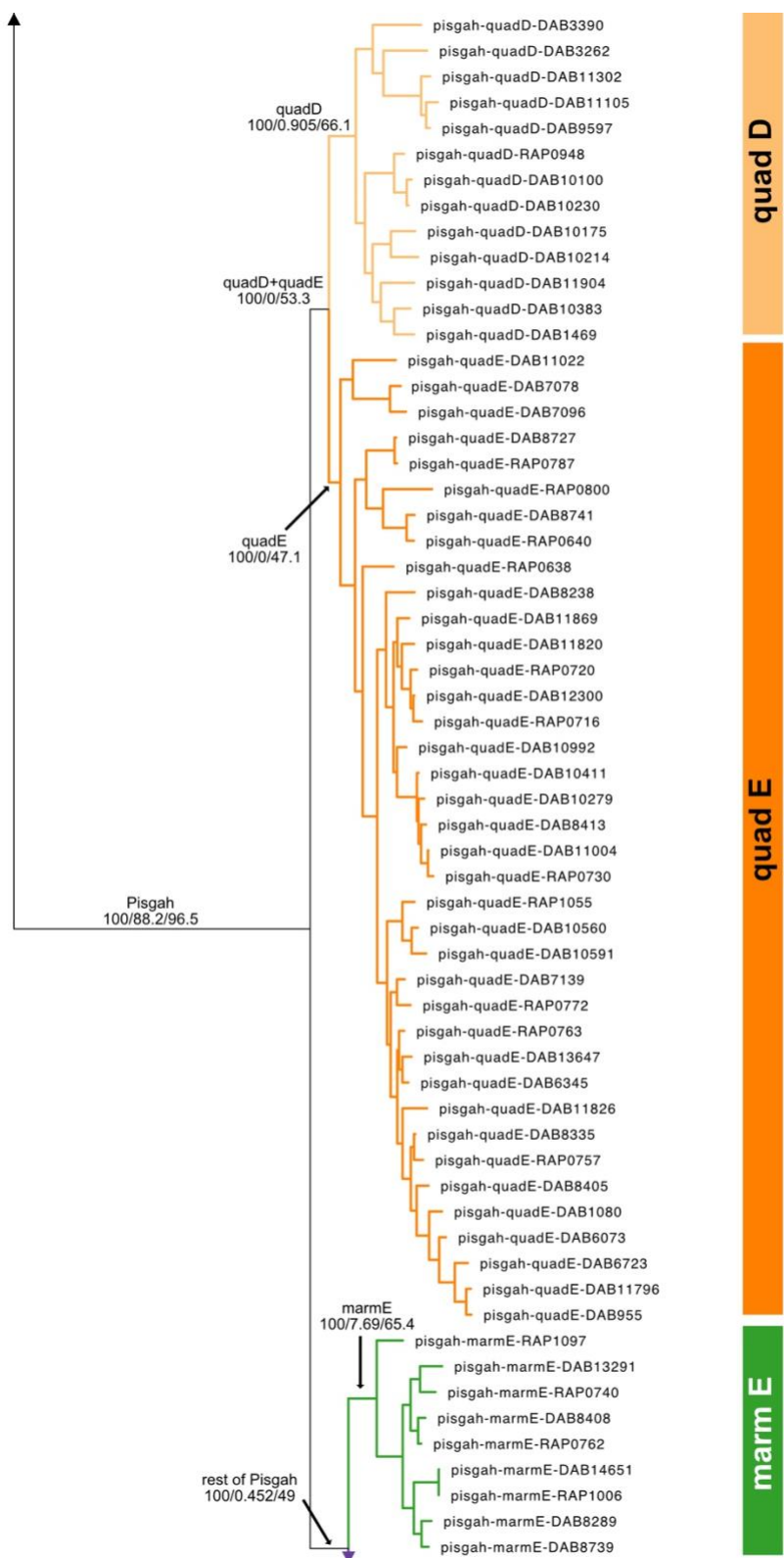


Figure 3.7 (continued)

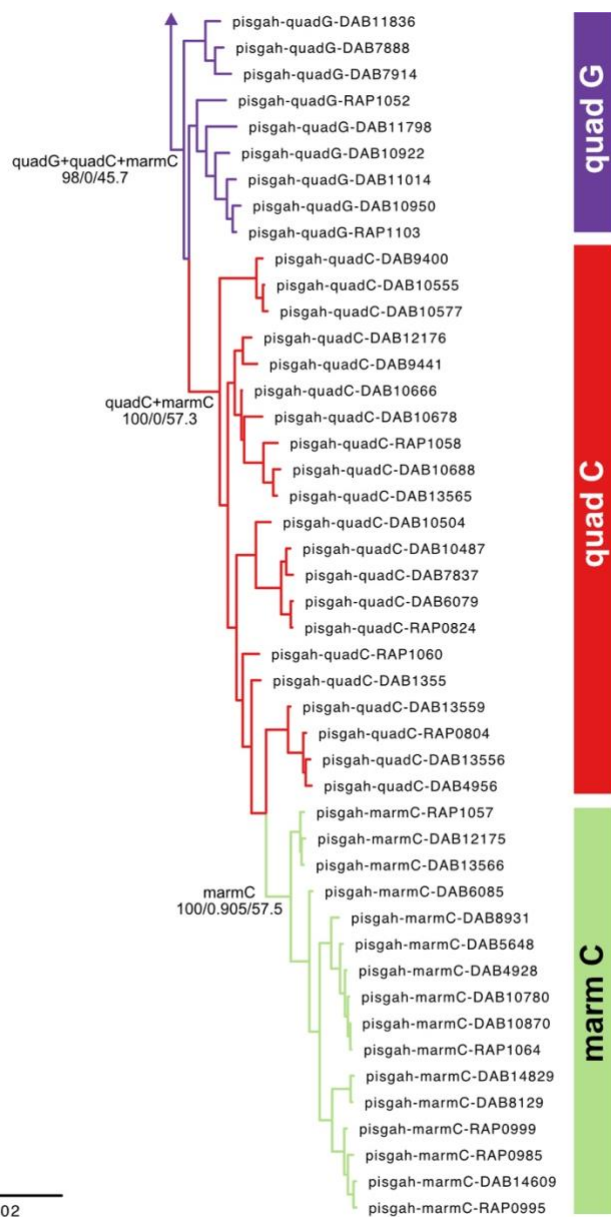


Figure 3.7 (continued)

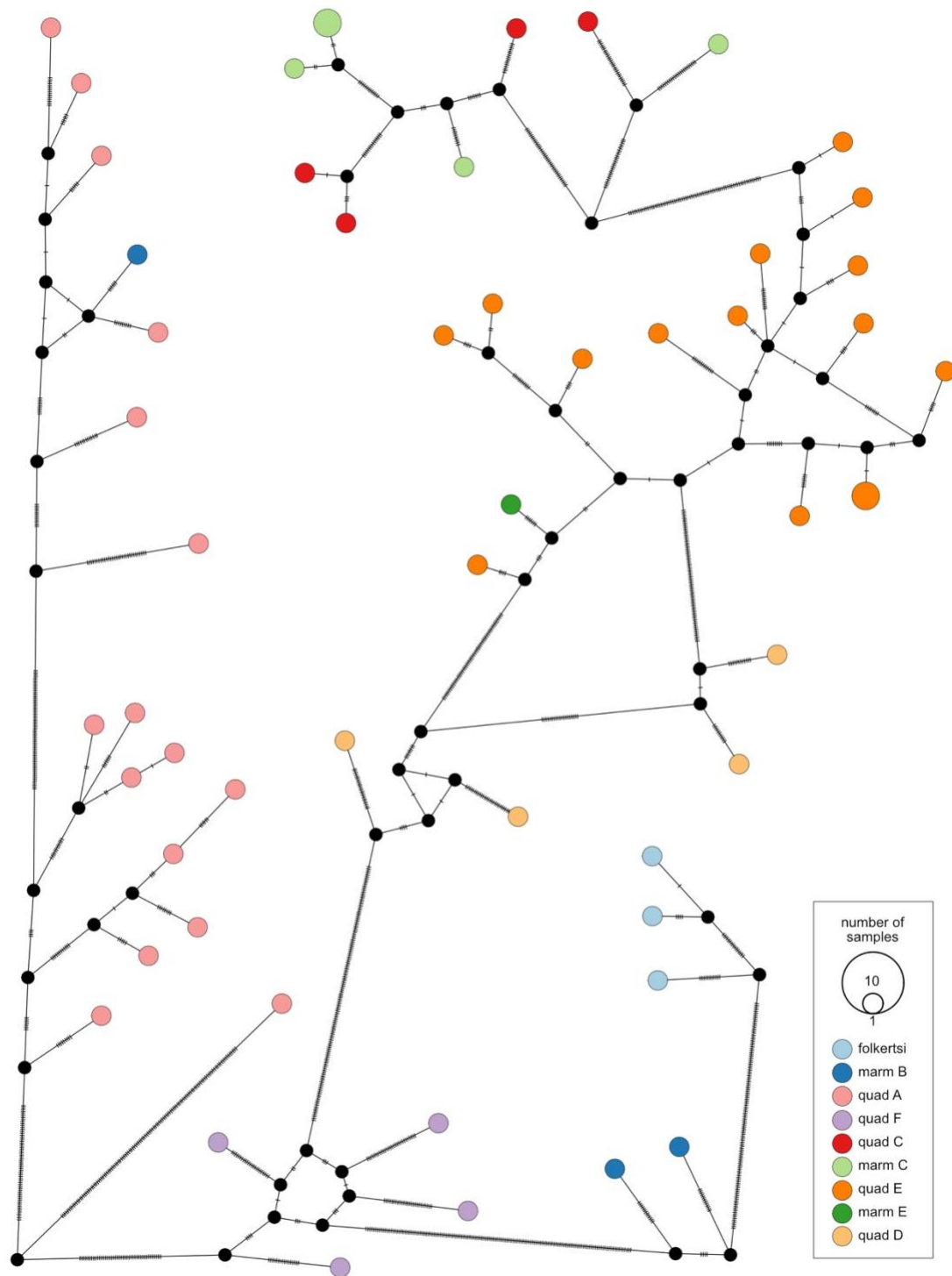


Figure 3.8. TCS haplotype network for COI. Tick marks represent one-step mutations. Each vertex represents a haplotype, colored by mitochondrial clade. Vertices sized by number of individuals with that haplotype. Colors follow *a priori* groups as per other figures.

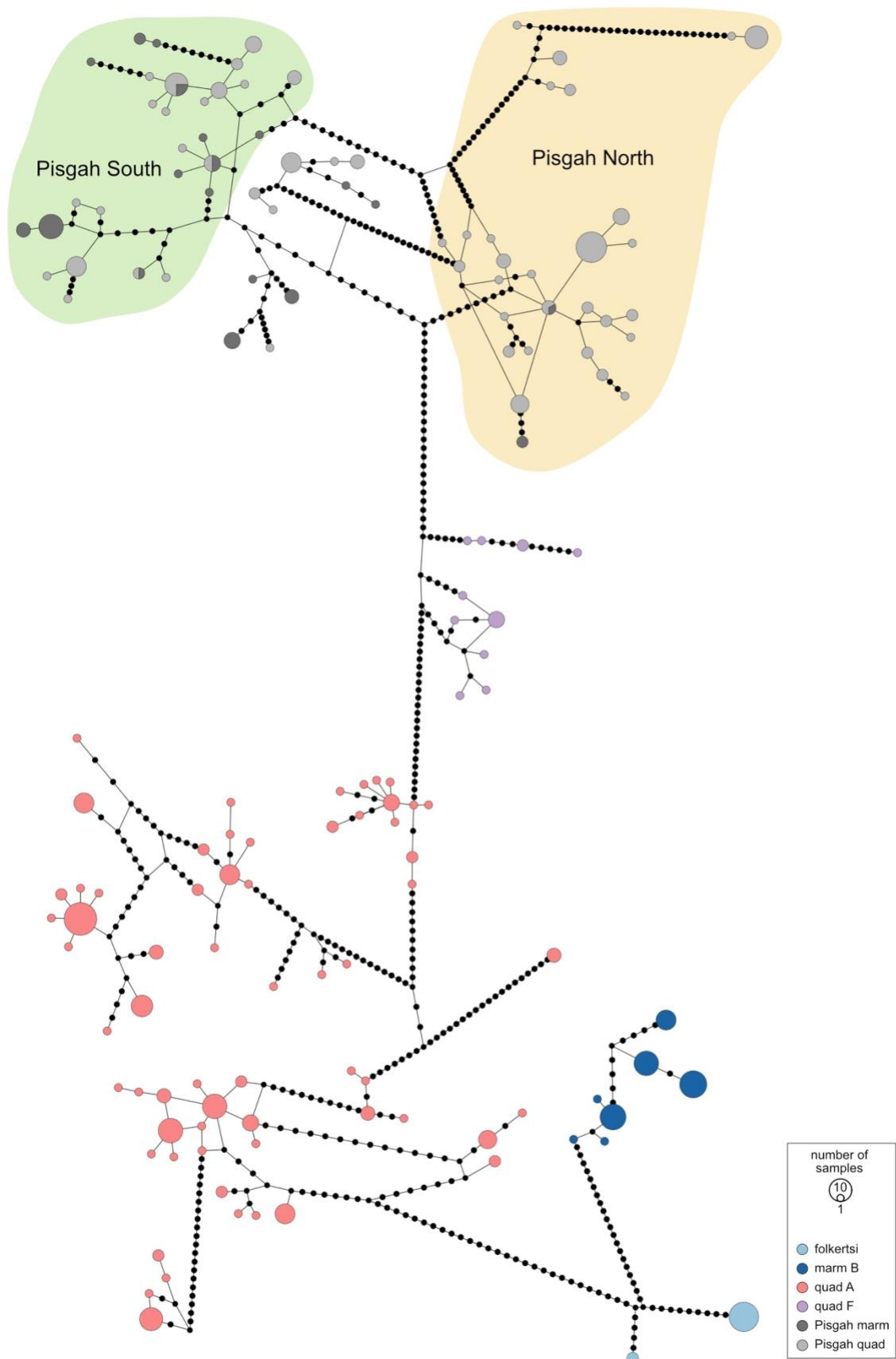


Figure 3.9. TCS haplotype network for CYTB. Each vertex represents a single haplotype and node radius represents the number of individuals associated with that haplotype. Nantahala haplotypes are colored by clade, while Pisgah are colored by phenotype. Pisgah haplotypes highlighted in yellow represent individuals that overlap geographically with the northern nuclear clades (quadE and quadD), green overlaps with southern clades (marm/quadC), and unhighlighted individuals overlap with the centrally located marmE and quadG. Solid black circles represent one-step mutations. Note: Edges that cross do not represent potential reticulations but alternative ways the haplotypes could be connected.

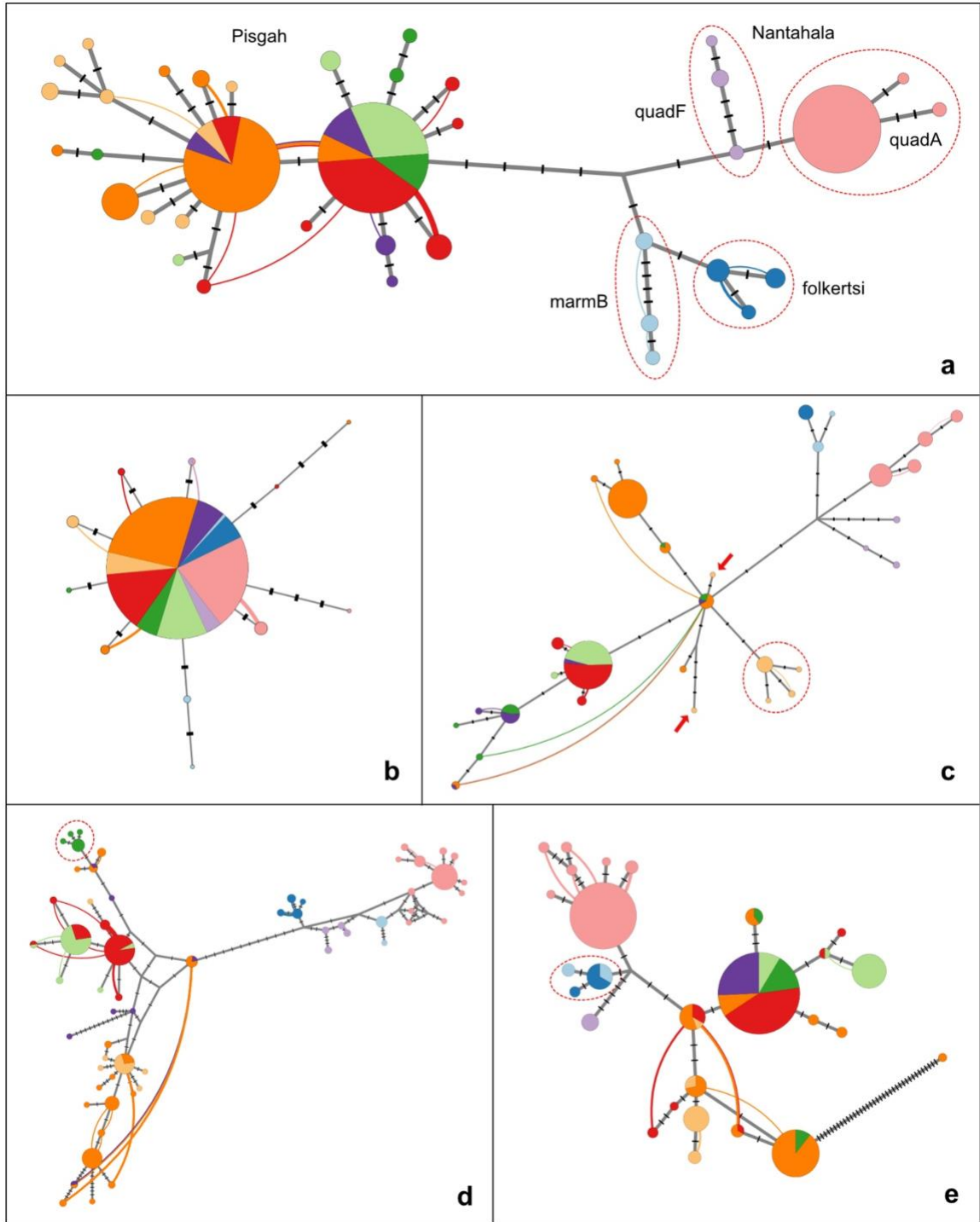


Figure 3.10. Example haplowebs for five nuclear loci. Grey lines connect haplotypes and colored lines represent curves connecting the two haplotypes found in heterozygous individuals. Black dashes indicate number of mutations between haplotypes. Radius of circles and width of lines represent relative number of individuals within each haplotype, while colors follow previous figures. a) Pattern typical for the majority of loci, with extensive interconnection between groups in Pisgah, while the four clades in Nantahala each have a set of exclusive haplotypes; b) individuals within Pisgah and Nantahala share a single haplotype; c) individuals with quadD-specific haplotypes form a cluster (circled), while other individuals in the group have disjunct haplotypes (red arrows); d) all individuals in marmE share one cluster of haplotypes exclusive to the group (circled); e) most individuals in folkertsi and marmB share the same haplotype.

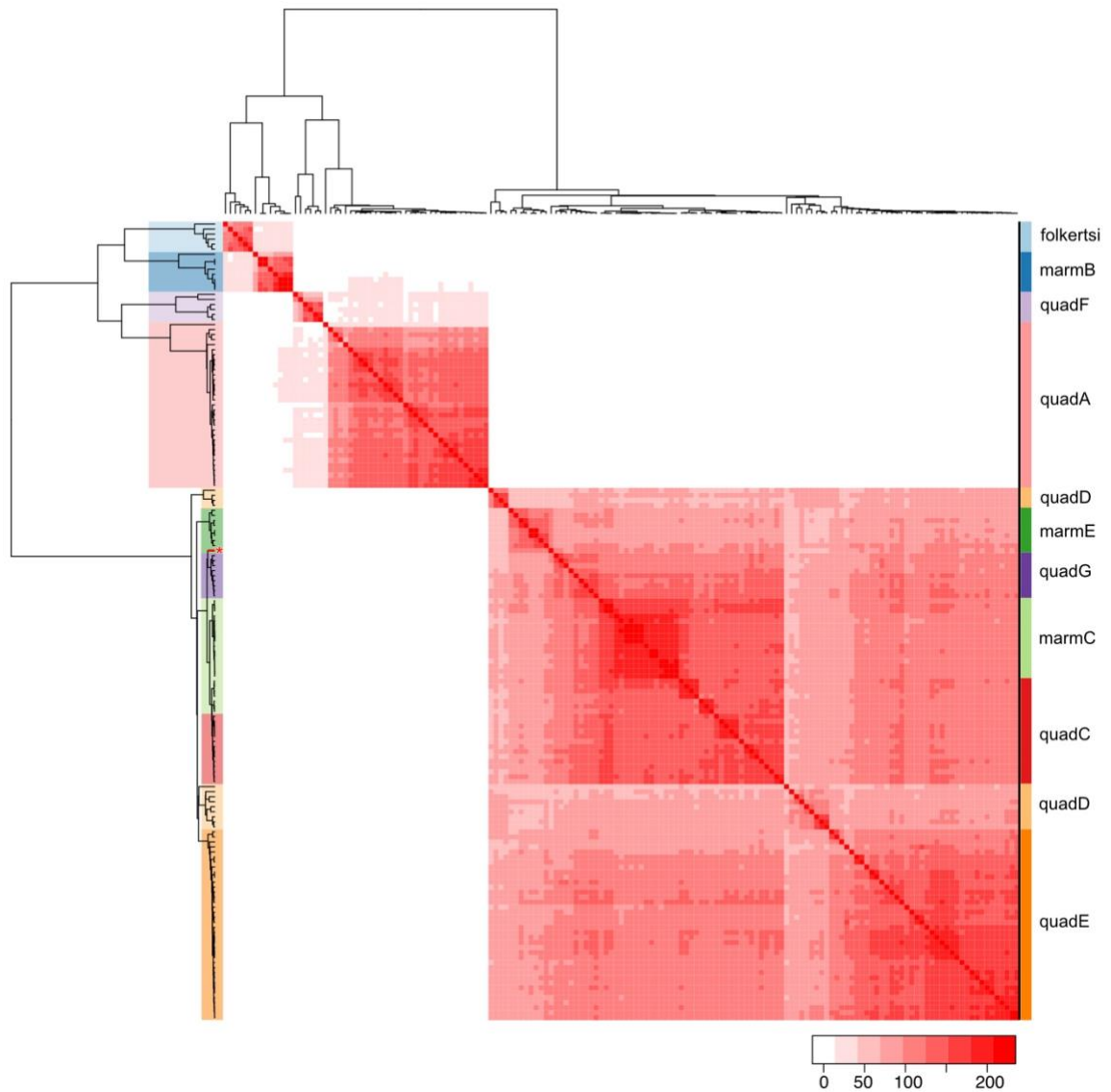


Figure 3.11. Conspecificity matrix analysis using the sum of conspecific sites. Individuals are grouped using UPGMA agglomeration. White squares have no loci supporting the grouping and darkest red squares have highest conspecificity (i.e., highest number of loci supporting the grouping). Colors highlighting portions of the phylogram correspond to groups as described in Pyron22. Note that quadD is separated into two groups and one individual in marmE groups with quadG (branch highlighted in red and asterisked on left phylogram).

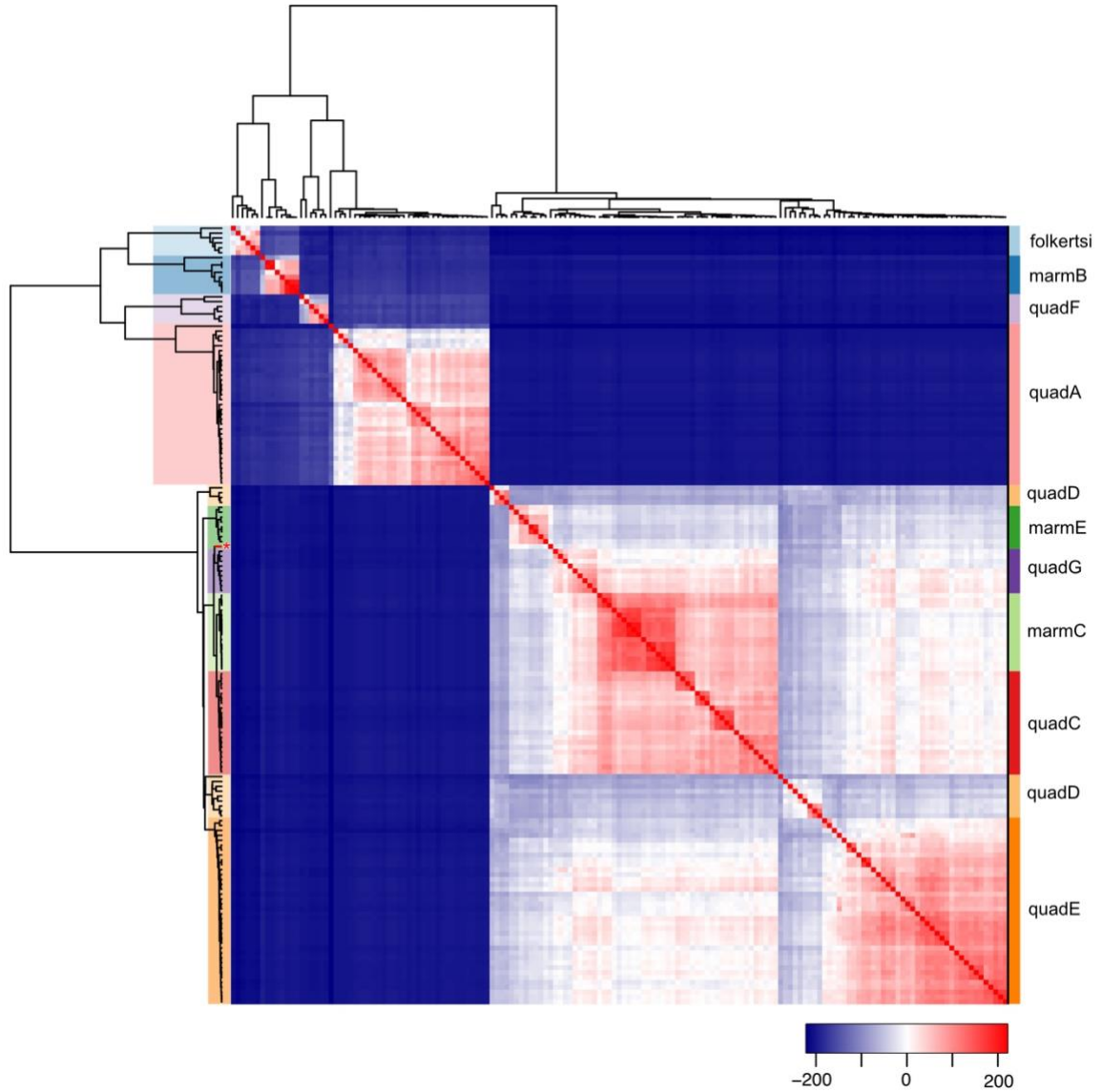


Figure 3.12. Results of conspecificity matrix analysis using the sum of conspecific loci minus the sum of heterospecific loci. Darkest blue squares indicate highest heterospecificity (i.e., least support for grouping), white squares have equal numbers of loci supporting for and against grouping, and darkest red squares have highest conspecificity (i.e., highest support for grouping). Note that quadD is separated into two groups and one individual in marmE groups with quadG (branch highlighted in red and asterisked on left phylogram)

CHAPTER 4. ECOLOGICAL NICHE MODELING PROVIDES INSIGHT INTO THE DISTRIBUTION OF GENETIC AND PHENOTYPIC DIVERGENCE

4.1 Abstract

Ecological data at different spatial and temporal scales can provide context for how evolutionary processes have influenced genetic divergence. Here, I used ecological niche models (ENMs) to examine how genetic and phenotypic diversity in the *Desmognathus quadramaculatus* salamander species complex was influenced by climate change and anthropogenic land use in the past, and how these factors may impact the complex in the future. I found that although the quadramaculatus and marmoratus phenotypes were nearly indistinguishable in niche space in the present day, they were predicted to occupy different geographic areas in the past and future, indicating that ecology may play a role in the distribution of phenotypes. Portions of the study area currently occupied by the Pisgah clade likely had little to no suitable habitat during the Last Glacial Maximum (LGM, ~22 kya), while portions occupied by the Nantahala clade had high habitat suitability from the LGM to the present, which aligns with the higher genetic divergence seen in the latter clade. Anthropogenic land use changes reduced habitat availability and likely contributed to the current distribution of individuals of all clades, which are largely limited to forest stands >40 years old that have seen little net forest loss over time. However, contemporary population structure likely reflects past climate change, rather than these recent anthropogenic land use changes. Future predictions indicate that climate change over the next 50 years may have less impact than changes in land use, such as increased development and loss of forest cover.

4.2 Introduction

Contemporary patterns of genetic variation are shaped by evolutionary processes acting across space and time (Araújo et al. 2008; Lowe et al. 2017). Gene flow and genetic drift are influenced by the landscape, with most species exhibiting some degree of isolation by distance as geographic distance exceeds dispersal capabilities (Aguillon et al. 2017; Cayuela et al. 2018). Population structure and lineage divergence can be influenced by physical barriers that prevent movement across the landscape or ecological barriers that cause species to alter their dispersal patterns to follow suitable habitat (Schluter 2001; Soltis et al. 2006; Sexton et al. 2014; Stroud and Losos 2016). Physical and ecological barriers can also change over time, such as when climate change renders previously occupied habitat unsuitable (Walther et al. 2002). Furthermore, gene flow and genetic drift can be relatively slow to act, leading to a delay between changes in population connectivity and measurable differences in genetic distance (Epps and Keyghobadi 2015; Maigret et al. 2020). Hence ecological data can provide insight into the processes that lead to genetic variation across the landscape at different time scales.

Ecological niche models (ENMs) provide a means to formulate and test hypotheses about how contemporary patterns of genetic variation arose (Peterson 2006; Richards et al. 2007; Warren 2012). Habitat suitability scores generated by ENMs can aid in determining the niche limits under which a species operates and provide insight into optimal habitat conditions (Soberón 2010; Nagaraju et al. 2013; Lee-Yaw et al. 2016). ENMs can be projected into the past to determine how habitat suitability has changed over time (Knowles et al. 2007; Smith et al. 2013), thus providing complementary evidence to assess whether population structure, lineage divergence, and taxonomic

boundaries inferred by genetic data are biologically plausible (Alvarado-Serrano and Knowles 2014; Pelletier et al. 2015; Dagnino et al. 2017). They can also be used to predict habitat availability into the future, giving insight into how climate change may alter a species' range, access to resources, and connectivity between populations (Searcy and Shaffer 2016; Ikeda et al. 2017; Micheletti and Storfer 2017).

Mountains provide an ideal landscape to study how spatial and temporal heterogeneity can impact genetic variation in a changing climate (Parmesan 2006; La Sorte and Jetz 2010; McCain and Colwell 2011; Condamine et al. 2018). For example, the Quaternary Period (2.6 mya to present) has been defined by rapid oscillations between warming and cooling periods, leading to geographic shifts in the availability of suitable habitat (Hewitt 2000; Davis and Shaw 2001; Nogués-Bravo et al. 2007). During warmer periods, species adapted to cool temperatures shifted to higher elevations as habitat suitability contracted, and unsuitable habitat in lowland areas served as a barrier to dispersal, isolating populations on individual peaks ("sky islands") and driving genetic differentiation (Crespi et al. 2003; Hedin et al. 2015; Hughes and Atchison 2015; Zhang et al. 2019). Conversely, suitability in lowland areas increased during cooler periods, when habitable lowland areas served as corridors for movement, allowing species to shift laterally across the landscape (Hewitt 2011; Thesing et al. 2016; Allen et al. 2021). Hence, even in areas free from glaciation, the current distribution and population structure of montane species reflect this history of alternating fragmentation and gene flow.

Along with these historical oscillations in climate, some mountain landscapes have also undergone recent changes due to human activity. For example, southern

Appalachia is renowned as a biodiversity hotspot (Dobson et al. 2008; Hodkinson 2010; Milanovich et al. 2010; Niemiller and Zigler 2013; Gilliam 2016), but its outward appearance as a seemingly pristine landscape belies a long history of land use changes that have altered the landscape extensively over time (Fraterrigo et al. 2005; Kuhman et al. 2011; Flatley et al. 2013; Woodbridge and Dovciak 2022). Portions of the region—such as the Asheville Basin—were likely used for agriculture long before European colonization (Delcourt et al. 2004; Gragson and Bolstad 2006). Post-colonization, agriculture expanded and large swathes of the forest were clear-cut for timber. Only a few small areas of primary (“old growth”) forest remain, with much of the area now covered with a patchwork of secondary regrowth interspersed with agricultural and urbanized areas (Turner et al. 2003; Steyaert and Knox 2008). Previous studies have shown that anthropogenic land use changes can negatively affect species by reducing available habitat and fragmenting existing populations, reducing gene flow and genetic diversity over time (Haddad et al. 2015; Newbold et al. 2015). However, land use is often excluded in ENMs and, to my knowledge, no previous studies have looked at how land use changes in southern Appalachia may have impacted population structure or genetic diversity in that region.

One of the most ubiquitous components of the southern Appalachian ecosystem are salamanders, particularly members of the *Desmognathus quadramaculatus* species complex, which show complicated patterns of genetic and phenotypic differentiation. The complex houses three distinct phenotypes which were originally described as separate species (*Desmognathus quadramaculatus*, *D. marmoratus*, and *D. folkertsi*), but previous studies disagree on whether phenotype is an indicator of genetic divergence (Chapter 3;

Jones and Weisrock 2018; Pyron et al. 2020, 2022; Pyron and Beamer 2022). The nominal species *D. folkertsii* and its dwarf phenotype appear to be genetically and morphologically congruent, but the *quadramaculatus* and *marmoratus* phenotypes may have evolved multiple times in two divergent lineages. While climate is suggested to have played a role in divergence between the two broader lineages (Jones and Weisrock 2018), it is unknown how changes in habitat availability have influenced genetic and phenotypic divergence in this group.

Hence the purpose of this study was to identify how climatic and anthropogenic changes have affected habitat availability over time to better understand how contemporary patterns of genetic variation arose in the *D. quadramaculatus* complex. To that end, I first used ENMs to determine whether there is a difference in niche space use between genetic clades and phenotypes. I then projected the ENMs in time slices back to the Last Glacial Maximum (LGM, 22 kya) to assess how changes in habitat suitability inform patterns of genetic variation seen today. I also assessed whether anthropogenic changes to land use since European colonization resulted in increased habitat fragmentation or instability. Finally, I projected ENMs into the future to see how connectivity might change between populations under future climate and land use change scenarios.

4.3 Methods

4.3.1 Study area and taxa

The Nantahala genetic clade is confined to areas south and west of the Asheville Basin and does not appear to extend beyond the eastern boundary of Great Smoky Mountains National Park. The Pisgah clade overlaps with portions of Nantahala in the south (but does not extend into Georgia and South Carolina, like the Nantahala clade) and above the Asheville Basin into southwestern Virginia. There are currently no studies using genome-wide markers that include individuals in West Virginia, but based on known patterns of genetic diversity it is probable any individuals north of North Carolina would be in the Pisgah clade and are therefore treated as such for the purposes of this study. The *marmoratus* and *quadramaculatus* phenotypes are found in both genetic clades, but *folkertsii* is only found in Nantahala.

Most of the Blue Ridge ecoregion is covered in Appalachian Oak forests, but widespread deforestation occurred in the mid-1800s to early 1900s (Steyaert and Knox 2008). Large swathes of the region are now protected, with Pisgah National Forest (2,075 km²) established in 1916, Nantahala (2,150 km²) and Cherokee (2,653 km²) National Forests in 1920, and the Great Smoky Mountains National Park (2,114 km²) in 1934. The region also contains several areas that are extensively urbanized, such as Asheville, North Carolina, and major roads run through potentially suitable habitat.

4.3.2 Occurrence and background data

I built separate ecological niche models (ENMs) for each of the three phenotypes and both genetic clades. For the ENMs comparing phenotypes, I downloaded occurrence data for *D. quadramaculatus*, *D. marmoratus*, and *D. folkertsi* from the Global Biodiversity Information Facility (GBIF), limiting observations to the last 30 years (1992-2022). I removed duplicate coordinates, records with no specific verbatim locality (i.e., sites that did not have the name of a stream or specific part of the road where the observation occurred), records with rounded coordinates, and records labeled as having increased coordinate uncertainty. To reduce outliers, I checked coordinates visually in QGIS v3.16 to ensure that they matched the verbatim description and removed observations outside of the ecoregions where the species are known to occur (i.e., Piedmont, Ridge and Valley, Blue Ridge, Central Appalachians; Figure 4.1). Nantahala and Pisgah can only be differentiated genetically, so occurrence data for those groups were taken from studies where both location data and clade were available (i.e., Jackson 2005; Jones and Weisrock 2018; Beamer and Lamb 2020; Pyron et al. 2022).

Since true absences are usually unknowable, ENMs are generally built with presence-only data where occurrences are compared against a set of background points. However, the selection of background data can introduce biases into the models. One common bias is that most occurrence data is collected in protected areas (Bowler et al. 2022), and near roads and other easily accessible areas (Kadmon et al. 2004; McCarthy et al. 2012). To account for this type of spatial bias, I created one set of background points drawn from GBIF observations of other plethodontid salamanders (genera *Desmognathus* and *Plethodon*) for the same range of years. I also created a second set of background

points designed to systematically sample the available habitat using a 5 km grid across the study area (Fourcade et al. 2014). Using the combined range of all the phenotypes to estimate the background also allows the model to project habitat suitability for each phenotype to any portion of the range. Replicate ENMs were built for each phenotype and clade using the same two sets of background points (spatial bias, SB, and gridded, GR; Table 4.1). All sets of occurrence and background points were thinned to 5 km using the *thin* function in R package spThin to reduce spatial autocorrelation, using 100 repetitions to select the maximum number of locations (Aiello-Lammens et al. 2015).

4.3.3 Environmental data

I used raster data sets of the standard 19 bioclimatic variables at 30 arc-seconds spatial resolution (1 km^2) to create baseline models for the ENMs from WorldClim v2.1 (Fick and Hijmans 2017) and CHELSA v1.2 (Karger et al. 2017). WorldClim is the most commonly used bioclimatic data set for species distribution modeling, but CHELSA may perform better in mountainous regions (Bobrowski et al. 2021). Hence I created ENMs with both data sets to assess their performance in southern Appalachia. While the algorithm used to build the models (MAXENT, see below) should theoretically be able to account for collinearity among variables, there is still disagreement over whether highly correlated variables should be removed from ENMs (Feng et al. 2019). I built each model using a full dataset with all 19 bioclimatic variables and a reduced data set where only variables with <0.75 correlation were included, for both WorldClim and CHELSA (Table 4.2). I also compared the full WorldClim and CHELSA data sets using Multivariate Environmental Similarity Surfaces (MESS) to

determine whether they predicted different areas of climatic similarity across the study area (Elith et al. 2010).

Since land use and vegetation data can improve the prediction of habitat suitability for ENMs (Regos et al. 2019), I tested models with two additional environmental data sets. The first was a land use and land cover (LULC) data set available for historical and future projections (Sohl et al. 2018). I reclassified the original 16 categorical variables into four categories likely to be biologically meaningful to the study taxa: 1) developed, 2) forest, 3) pasture, and 4) crop (Table 4.3). Open water and ice were considered uninhabitable and excluded from the models (i.e., set as NA). To provide additional information on forest quality, I also included data for historical forest stand age (Sohl et al. 2018). For LULC and forest stand age data, I used the year 2000 as the baseline because it fell within the range of dates used as the average historical baselines by WorldClim v2.1 (1970-2000) and CHELSA v1.2 (1979-2013). All rasters were reprojected to the same coordinate reference system (WGS84; EPSG:4326) and resampled to 1 km² using a nearest neighbor algorithm for categorical data and bilinear interpolation for continuous data.

Two variables that could also be important to the *quadramaculatus* complex that were not included in the final models are altitude and proximity to streams. Altitude was excluded from models because it had high collinearity with several temperature variables and occurrence at different altitudes probably reflects the ecological attributes found at that altitude (i.e., lower temperatures) rather than the exact altitude itself.

Proximity to streams was excluded because the minimum resolution of the models was 1

km² due to the minimum available resolution for climate data and almost all background points were within 1 km of a stream (data not included).

4.3.4 Model calibration and evaluation

I built ENMs using a Maximum Entropy machine-learning model (MAXENT) which has proven to be the most accurate for niche modeling when compared against different methods across a variety of species and regions (Ashraf et al. 2017; Kjeldsen et al. 2019; Venne and Currie 2021; though see Escobar et al. 2018). I used the R package ENMeval to tune MAXENT models using linear, quadratic, and linear-quadratic feature classes, and regularization multipliers between 0.5-6.0 at 0.5 increments (Merow et al. 2013; Muscarella et al. 2014). I did not include the hinge feature class because it can produce biologically unrealistic cut-offs for environmental variables (Guevara et al. 2018). To account for spatial autocorrelation, occurrence and background points were split into four equal groups partitioned spatially by latitude and longitude using the block method, which produces more realistic estimates than other methods (Santini et al. 2021).

The model with the lowest AICc score was considered the “best” model (Galante et al. 2017). Since $\Delta AICc$ values below two are not considered meaningfully different, models with an $\Delta AICc < 2$ were further filtered sequentially by lowest average test omission rate and then highest average validation area under the curve (AUC) to determine the best model. Each of the models was compared to a null model built using the same model settings with 1000 iterations of randomized data to determine if models outperformed the null for AUC and the Continuous Boyce Index (CBI; Hirzel et al. 2006; Bohl et al. 2019).

I used permutation importance to rank variables for each model to determine which models were most important for habitat suitability (Barbet-Massin and Jetz 2014; Searcy and Shaffer 2016; Smith and Santos 2020). I also examined model response curves to determine whether suitability was positively or negatively correlated with variables that had high permutation importance (Sillero et al. 2021). Rather than applying a threshold to reduce presence/absence to a binary—which reduces the predictive ability of the model—habitat suitability is presented as a probability for all habitat suitability comparisons (Liu et al. 2016; Santini et al. 2021).

I used two niche identity tests (also known as niche equivalency tests) to determine whether each phenotype and clade occupied different niche space. The first was a geographic identity test, which pools occurrence points and randomly reassigns them to two test groups. ENMs are then built for each resampled set of occurrence points and niche similarity indices (Schoener's D and Warren's I) are calculated for 100 pseudoreplicates. This produces a distribution of scores for D and I under the assumption that the niche space used by the two groups is equivalent (Warren et al. 2008, 2010). If the empirical value of D or I is significantly lower than the distribution of scores from pseudoreplicates, then it indicates that the niches are not equivalent. The second is an environmental identity test which uses a principal component analysis (PCA) to transform the environmental variables into two axes, with each axis representing a set of correlated variables. Niche space for each ENM is represented by the kernel density of occurrence points that fall within a grid of PCA scores found across the two species. The distribution of D and I is calculated and interpreted the same way as for the geographic identity test (Broennimann et al. 2012; Guisan et al. 2014; di Cola et al. 2017).

4.3.5 Prediction of past and future habitat suitability

To determine how habitat suitability may have changed over time, I built ENMs for every 1000 years from the present (1900-1990 CE) to the Last Glacial Maximum (20,100-20,001 BCE) using the CHELSA-Trace21k v1 data set (Karger et al. 2021), which uses CHELSA v1.2 climate data averaged over 100 year time steps for the last 22,000 years. To match the baseline and past data temporally, ENMs were projected using 1900-1990 CE as the baseline, and best models were selected as described above. Note that LULC and vegetation were not included in these models because they are not available for this time period. Habitat suitability was predicted using the *predictMaxNet* function in the R package *enmSdm*, with variables clamped to the model training data (Morelli et al. 2020). Predictions were output as probabilities on a complementary log-log scale to match suitability predictions for baseline models.

For future predictions, I used CHELSA v1.2 CMIP5 bioclimatic variables (Karger et al. 2017, 2020) to extrapolate habitat suitability from the CHELSA+LULC baseline average (1979-2013) to two future time periods: 2041-2060 and 2061-2080 (hereafter, the 2050 and 2070 models, respectively). To account for uncertainties between General Circulation Models (GCMs; Thuiller 2004; Varela et al. 2015), I used three GCMs that had relatively low interdependence and high skill weight for North America, ACCESS1-0, CCSM4, and CESM1-CAM5 (Sanderson et al. 2015; Eyring et al. 2019), at two representative concentration pathways (RCP) that assumed carbon emissions would either stabilize (RCP4.5) or continue to increase (RCP8.5). For LULC and forest stand age data, I used the storylines from the Intergovernmental Panel on Climate Change Special Report on Emissions Scenarios (SRES) that best matched the climate projections

(i.e., the SRES B1 scenario with RCP4.5 climate models and the SRES A2 scenario with RCP8.5; Riahi et al. 2017). Future model predictions were performed as for the LGM predictions described above. Final suitability rasters were summarized across future climate projections using the Weighted Sum Raster Overlay function in QGIS, with each prediction given equal weight.

4.3.6 Changes in land use between 1650-2020

I calculated changes in land use and percent forest cover between 1650-2000 to determine how anthropogenic land use may have impacted the study area since European colonization. I used land use (LU) rasters, available as yearly boolean categorical data where each 1 km² raster cell was assigned to either forest, crop, pasture, or developed. Additionally, I used percent forest cover, available as the percentage of each 1 km² raster cell covered in forest (Li et al. 2022). I calculated the year-on-year stability for LU and forest cover using the R package *climateStability*, where stability is the inverse of the mean standard deviation among years (Owens and Guralnick 2019). To reduce the impact of small percentage changes, forest cover was reclassified into four coverage categories from low to high (0-25%, >25-50%, >50-75%, >75%) prior to calculating stability. I also assessed the impact of deforestation by calculating change in percent forest cover at six snapshots in time (1650, 1750, 1850, 1900, 1950, and 2000). Since forest cover was calculated as a percentage across each 1 km² raster cell, it can serve as a rough proxy for forest fragmentation within an area. Change in forest cover between time periods was calculated as mean cover across all cells using the *cellStats* function from the R package *raster*. All rasters were cropped to the study area prior to performing calculations.

Since there is a delay before a change in gene flow is reflected in genetic distance, current genetic distances may reflect habitat availability in the past rather than current conditions (Epps and Keyghobadi 2015; Maigret et al. 2020). Thus, to determine how landscape changes may have impacted gene flow within the Nantahala and Pisgah clades I conducted a landscape resistance analysis to compare pairwise genetic distance between individuals (D_{PS} ; see Chapter 3 for details) to resistance distance at the same six time periods used above. Resistance analyses are based on circuit theory and estimate landscape connectivity by calculating all possible routes between two or more areas where a species is known to occur (Hanks and Hooten 2013). Hence, resistance surfaces (or their inverse, conductance surfaces) that accurately reflect habitat corridors suitable for dispersal should produce distances that are more closely correlated with genetic distance than straight-line geographic distance (McRae 2006; McRae and Beier 2007).

I created conductance surfaces using the assumption that a higher percentage of forest cover would serve as better habitat corridors and therefore provide greater conductance to movement. I used the R package *gdistance* to convert each raster to a transition layer, allowing for movement between cells in 16 directions, and using mean forest cover percentage as the transition function (van Etten 2017). I then applied geocorrections to correct map distortions that could affect resistance distances. I calculated resistance distance using the average random-walk time between nodes calculated across all possible routes (i.e., commute distance), taking into account the conductance weights along the routes. Hence, resistance distances between two individuals would be smaller if they were separated by areas of high forest cover and larger if they were separated by areas of low forest cover. I tested for significant

correlations between genetic distance and resistance distance for Nantahala and Pisgah separately using Mantel tests with 10,000 permutations using the *ecodist* package (Goslee and Urban 2007; Goslee 2010). For comparison, I also calculated the correlation between genetic distance and Euclidean geographic distance (i.e., isolation by distance), topographic distance (i.e., geographic distance weighted by altitude), and resistance distance against a null model where all pixels were set to 1 (full conductance).

4.4 Results

4.4.1 Ecological niche models

CHELSEA data had finer-scale heterogeneity across altitudes for annual mean temperature (BIO1) and precipitation (BIO12) compared to WorldClim (Figure 4.2). The MESS analyses for CHELSEA showed that a higher percentage of the study area had similar conditions to those found at occurrence points and more heterogeneous climate conditions compared to WorldClim, especially in the north and south of the study area (Figure 4.3). Only six of the models built with SB backgrounds performed better than null models, while all but one model built with GR backgrounds had significantly higher values of AUC and/or CBI compared to the null models (Table 4.4). Models performed similarly regardless of whether the full set of 19 bioclimatic variables or only reduced correlation variables was used. Overall, the CHELSEA+GR models had significantly higher AUC and CBI than the null for all the groups tested, while the WorldClim+GR models varied. Models with and without LULC performed similarly based on AUC and CBI, but areas of high suitability were patchier in models that included LULC (Figures 4.4 & 4.5). Categorical LULC did not have high permutation importance values for any

model, but forest stand age had high permutation importance for *marmoratus* and *quadramaculatus* (Table 4.5).

The identity tests indicated that all groups had significantly different use of geographic and environmental space, except for *marmoratus* versus *quadramaculatus*, which did not have significantly different use of environmental space (Table 4.6) and had nearly complete (0.92) niche overlap in environmental space (Figure 4.6).

4.4.2 Past habitat suitability predictions (LGM to present)

Areas of high habitat suitability for *folkertsi* were patchy and confined mainly below the Asheville basin 22 kya, then suitability contracted to its lowest point during the Younger Dryas (12-13 kya). After the Younger Dryas, swathes of high suitability appeared below *folkertsi*'s current range and spread northwards to form the current pattern of suitability seen at 0 ya (i.e., the average climate over the last 1000 years). For *quadramaculatus*, areas of high suitability before the Younger Dryas were confined primarily to lowland areas. After the Younger Dryas, areas of high suitability began to appear in the southern portion of the range in both high and lowland areas, after which it contracted steadily over time from lowland areas. For *marmoratus*, almost the entire southern portion of the study area had high suitability 22 kya. This swath of high suitability moved steadily upward and encompassed all of the study area from 17 kya to 14 kya. After the Younger Dryas, areas of high suitability contracted from lowland areas. Overall, habitat suitability has been trending downwards since the Younger Dryas for *marmoratus* and *quadramaculatus*, but peaked for *folkertsi* 1-4 kya and is now decreasing (Figure 4.7).

Nantahala and Pisgah show contrasting patterns of suitability over time, with Nantahala showing an overall contraction—and Pisgah an expansion—in suitability tracing from the LGM to 0 ya. For Nantahala, almost the entire study area had high suitability 22 kya, while suitability for Pisgah was low during the same time period. Around 13-14 kya, areas of high suitability for Nantahala began to decrease, starting in the northeast of the study area and contracting over time until it was limited largely to mountainous areas below the Asheville Basin. For Pisgah, the first patches of mid-level (40%) suitability appear around 13-14kya. Areas of high suitability appear around 10 kya, but the locations of high and low suitability areas were unstable until around 5 kya, when high suitability stabilized in areas similar to where they appear at 0 ya.

4.4.3 Future habitat suitability predictions

Summarized future climate projections for most models predicted that decreases in habitat suitability would be limited largely to lowland areas that members of the *quadramaculatus* complex do not currently occupy. Habitat suitability for *quadramaculatus* is predicted to increase by 2050, particularly in the northern portion of the range, followed by a slight decrease in the southern portions of the range by 2070. Areas of high suitability for *marmoratus* shift slightly northwards by 2050, followed by a contraction from lowland areas by 2070. Areas of high habitat suitability for *folkertsi* are predicted to increase across the southern portion of the range. For Nantahala, suitability increases above the Asheville Basin by 2050, then contracts from lowland areas and increases in higher elevations areas by 2070. For Pisgah, suitability decreases across most

areas by 2050, except for an increase in northern portions of the range, followed by a contraction of high suitability to higher elevations across the range by 2070 (Figure 4.8).

4.4.4 Changes in land use between 1650-2020

Percent forest cover decreased from 96.5% to 70.3% between 1650 and 2000 (Figure 4.9). The biggest drop (-13.5%) occurred between 1900 and 1950. Occurrences were primarily in areas that had little net forest loss over time (Figure 4.10) and where forest stand age was at least 50 years old (Figure 4.11). Despite widespread instability in percent forest cover over time, forest remained the most stable dominant land use category over most of the study area (Figure 4.12). In areas that were converted from forest to other land use categories, 8.7% were converted to urban development, 12.3% to pasture, and 0.2% to crops (Figure 4.13).

For both Nantahala and Pisgah, the correlation between genetic distance and resistance decreased between 1650 and 1900, rebounded slightly in 1950, and then decreased again in 2000 (Figure 4.14). There was a significant correlation between genetic distance and resistance distance for all time periods in Pisgah ($R = 0.51-0.50$, $p = 0.0001$), but only for 1650 ($R = 0.17$, $p = 0.023$), 1750 ($R = 0.18$, $p = 0.024$), and 1850 ($R = 0.16$, $p = 0.027$) in Nantahala. There was a significant correlation between genetic distance and geographic distance (Nantahala: $R = 0.15$, $p = 0.009$; Pisgah: $R = 0.52$, $p = 0.0001$), topographic distance (Nantahala: $R = 0.10$, $p = 0.014$; Pisgah: $R = 0.52$, $p = 0.0001$), and resistance distance for the null model (Nantahala: $R = 0.13$, $p = 0.016$; Pisgah: $R = 0.55$, $p = 0.0001$).

4.5 Discussion

4.5.1 Impact of variable selection on ENMs

There were substantial differences in the ENMs built using WorldClim v2.1 and CHELSA v1.2 bioclimatic data sets. CHELSA consistently outperformed WorldClim, with ENMs built using CHELSA having higher positive CBIs than equivalent WorldClim models, indicating that the CHELSA models predicted presence with greater accuracy (Table 4.4). These differences may be due to the incorporation of small-scale orographic effects in CHELSA, the inclusion of which improves climate projections in topographically complex areas (Karger et al. 2017). Previous studies have shown that CHELSA outperforms WorldClim in the Himalayas (see Bobrowski et al. 2021 for a review), which is unsurprising considering that it was one of the mountain ranges explicitly validated during the development of CHELSA. However, the trend of increased accuracy appears to hold in southern Appalachia despite less elevational variation in the region. These results indicate that future studies building ENMs for species in southern Appalachia should consider using CHELSA as an adjunct or alternative to WorldClim.

CHELSA models with and without the inclusion of LULC data had similar CBI values, but models that included LULC had lower suitability overall, and areas of high suitability showed patchier distributions (Figures 4.4 & 4.5). Poor performance in models using the spatial bias (SB) background points—which already had limited power due to a lower number of training points compared to the grid background (Barbet-Massin et al. 2012)—may have been exacerbated by undersampling across LULC categories. MAXENT requires sufficient input from across different variables for accurate model training and almost all occurrence and background points fell within 1 km of forest (i.e.,

the minimum resolution of the model). Thus all models were limited in their predictive performance for areas in non-forest land use categories (i.e., developed, pasture, and crop) regardless of background due to the homogeneity of land use in the study area.

While categorical land use (e.g., developed versus forest) had low permutation importance in all models, the inclusion of forest stand age produced suitability estimates that were lower in developed areas compared to models lacking this variable. Models that included forest stand age had response curves that showed a strong positive correlation between forest stand age and habitat suitability, and occurrence in older forest stands was higher than expected based on habitat availability (Figure 4.11). However, the permutation importance of LULC varied widely between phenotypes and clades, and models including LULC still overestimated suitability in areas known to be unsuitable, like urban centers. Climate and LULC data may compete for importance in models since these two factors are not independent. For example, impervious surfaces in developed areas generally result in higher temperatures and altered precipitation regimes compared to surrounding areas due to the urban heat island effect (Steensen et al. 2022). LULC may have also been skewed by the reclassification of different types of development into a single category, which does not take into account tolerance to some forms of anthropogenic disturbance. A compromise approach to effectively denote areas that are unlikely to be suitable regardless of climate may be to model climate variables separately and then overlay land use onto suitability maps (Figure 4.8). Overall, these results highlight how much impact the choice of variables can have on ENMs and the limitations inherent in models trying to incorporate the effects of anthropogenic land use.

4.5.2 Historical habitat suitability

Ecological niche models for the Nantahala clade predicted that portions of the study area currently occupied by that clade had high suitability beginning in at least the Last Glacial Maximum (Figure 4.7). The distribution of land cover available in southern Appalachia during the LGM is still poorly understood. Older fossil pollen studies suggested that deciduous trees were extirpated from southern Appalachia during the LGM (Davis 1983; Delcourt and Delcourt 1988), but more recent work shows that deciduous trees and understory plants associated with modern mesic forests probably did occupy southern portions of the study area (hickory [*Carya* spp.], Bemmels and Dick 2018; sweet birch [*Betula lenta*], Thomson et al. 2015; red maple [*Acer rubrum*], McLachlan et al. 2005; American bellflower [*Campanulastrum Americanum*], Barnard-Kubow et al. 2015; wood lily [*Trillium cuneatum*], Gonzales et al. 2008. Furthermore, current evidence indicates that discontinuous permafrost may have extended as far south as the Appalachian Basin and areas near the border of Georgia and South Carolina (French and Millar 2014; Lindgren et al. 2016, 2018), which would have left most of the range of the Pisgah clade covered in permafrost and the range of the Nantahala clade largely free of permafrost (Figure 4.7).

Together, these results suggest that the northern portion of the study area may have contained limited and patchily distributed suitable habitat for *quadramaculatus* and *marmoratus* during the LGM, while large contiguous areas of suitable habitat would have been continuously available across the range of Nantahala beginning at least during the LGM. This supports the results of genetic studies showing higher levels of divergence between populations within Nantahala, since most areas of Pisgah north of the Asheville

Basin may not have been able to support salamander populations until relatively recently (~5000 ya). However, it also indicates that predictions of habitat suitability may be overestimated for Nantahala during the LGM in regions above the permafrost line. Furthermore, ENMs for present-day habitat suitability in Pisgah show low suitability in some currently occupied areas. Hence, it is likely that the genetic ENMs provide an incomplete picture of suitability.

All the phenotypes differ substantively in their responses to past (and future) climate predictions. This is surprising considering that the *marmoratus* and *quadramaculatus* phenotypes have nearly complete (92%) environmental niche overlap in the present day, indicating that their use of niche space is practically indiscernible under current climate conditions (Table 4.6, Figure 4.6). During the LGM, *marmoratus* had higher suitability over larger portions of the study area than *quadramaculatus*, and the ranges of the two phenotypes have converged over time to produce similar distributions only during the last 1000 years (Figure 4.7). Suitability for *folkertsi* has also changed substantially, with low overall suitability across the study area for most of the past, leading to range overlap with *marmoratus* and *quadramaculatus* in the most southern portions of Nantahala as suitability for *folkertsi* expanded northward. Furthermore, occurrence over the last 30 years aligns more closely with areas of highest suitability ~5000 ya rather than predicted suitability during the present day, or closer to the LGM, for all the phenotypes. This indicates that the ENMs either underestimated suitability in some parts of the study area, or that the current distribution of the phenotypes includes remnant populations isolated as suitability began to contract during the last 1000 years. While the unoccupied gaps between some populations lend weight to the latter scenario,

it is important to note that the occurrence data were largely drawn from citizen science sightings which are subject to bias, and thus these gaps could be artifacts of the data.

4.5.3 Effects of anthropogenic land use on population structure

Widespread loss of forest cover occurred in southern Appalachia between the late 1800s to early 1900s, followed by regrowth after the establishment of several National Forests and the Great Smoky Mountains National Park in the 1910s-1930s (Steyaert and Knox 2008). While members of the *quadramaculatus* complex occur primarily in areas where forest was the dominant land use category from the 1650s onwards, occurrence is not limited to areas that had the most stable percentage of forest cover over time. Instead, occurrence is concentrated in areas that show little to no *net* forest change over the last 300 years (Figure 4.10), primarily in stands that are at least 50 years old (Figure 4.11). This finding aligns with previous studies showing significant recovery of salamander species diversity within ~40 years post-logging (Petranka et al. 1994; Ash 1997). However, it may take over 100 years for previously logged forests to reach peak carrying capacity for salamanders, and most forest stands in the region consist of secondary growth less than 75 years old. Thus, while areas of southern Appalachia that were logged in the early 1900s may have recovered species diversity, population sizes for some species may still be lower than they were prior to anthropogenic disturbance.

Previous studies on the impact of logging on *quadramaculatus* have been mixed, with some studies showing that logging does not reduce abundance in streams, unlike other streamside and terrestrial salamanders (Stiven and Bruce 1988; Peterman et al. 2008), and others showing population crashes post-logging (Petranka et al. 1993, 1994;

Cecala et al. 2018). *Quadramaculatus* can persist in portions of streams without canopy cover—at least over short periods of time—but individuals experience a loss of body condition in canopy gaps, which may lead to higher mortality and local extirpation over the long term (Bliss and Cecala 2016). Canopy gaps can also significantly reduce stream-mediated dispersal in *quadramaculatus* (Cecala et al. 2014), potentially leading to population fragmentation. Hence populations that now occur in contiguous patches of forest are likely composed of members from smaller patches that were difficult to clear-cut, such as steep slopes, that then recolonized larger areas as the forest regrew. These results indicate that the *quadramaculatus* complex in southern Appalachia likely saw large population bottlenecks when areas were logged, followed by decreased gene flow between newly isolated populations until forest regrowth made habitats suitable for occupancy again. Fragmentation from logging took place after the *marmoratus* and *quadramaculatus* phenotypes had already lost suitable habitat due to climate change over the last 4000 years, suggesting that population sizes may have been trending downward for an extended period prior to large-scale anthropogenic disturbances.

A demographic bottleneck would be expected to leave a genetic signature of decreased genetic diversity and increased divergence between populations (Nadeau et al. 2017), neither of which were seen in the pairwise genetic distances and population structure analyses performed in Chapter 3. This is particularly noticeable in areas such as the Asheville Basin, which had most of its forest cover converted to other land use categories and probably continues to act as an ecological barrier to gene flow, yet there is no evidence of genetic divergence between populations on either side of this divide. In areas of suitable habitat, the *quadramaculatus* phenotype occurs at particularly high

population densities compared to other salamanders (Peterman et al. 2008), which may have allowed them to maintain high standing genetic variation despite extensive habitat loss, and thus recolonize the forest as it regrew without losing allelic diversity (Pabijan et al. 2020). Thus *quadramaculatus* may not have undergone a severe enough reduction in effective population size to overpower signals of population structure that were in place prior to logging.

Additionally, *quadramaculatus* were only relegated to isolated patches for relatively short periods of time, and thus may not have been isolated long enough to experience enough genetic drift to produce diagnosable markers of population divergence. Genetic distances are also most highly correlated with resistance distances from 1650, when forest cover was at its highest, indicating that habitat fragmentation did not lead to detectable changes in genetic distance. Hence while it can be difficult to determine whether diversification and population structure is driven by historical habitat availability or reflects ongoing gene flow (Samarasin et al. 2017; Vandergast et al. 2019), genetic structure in this species complex likely reflects long-term historical processes rather than the recent effects of logging.

4.5.4 Effects of anthropogenic change on habitat suitability

Montane species are generally projected to experience range contractions or aspect shifts in response to anthropogenic climate change (La Sorte and Jetz 2010; Feldmeier et al. 2020). Shifts in salamander ranges in southern Appalachia have been recorded within the last 50 years, but it is uncertain whether these shifts are due to climate change, forest regrowth, interspecies competition, or a combination of factors

(Walls 2009; Moskwik 2014; Grant et al. 2018). Some low elevation plethodontid species have already been shown to be at their thermal limits, with summer temperatures leading to metabolic depression (Bernardo and Spotila 2006), while others can acclimatize to hotter and drier conditions while maintaining energy balance (Riddell et al. 2018). In mesocosm experiments with *quadramaculatus* phenotypes, elevated temperatures resulted in decreased body condition for individuals (Hoffacker et al. 2018), as did lower stream flow under experimental drought conditions (Currinder et al. 2014).

However, most studies examining the effects of climate change in plethodontids assume that climate change in southern Appalachia will lead to hotter, drier conditions, which may not be entirely accurate. Compared to 1950, portions of southern Appalachia have recently experienced a cooling trend or little change in annual temperature, along with an increase in precipitation (Sayemuzzaman et al. 2015; Lesser and Fridley 2016; Eck et al. 2019). Depending on the climate model used, rises in annual mean temperature across southern Appalachia over the next 50 years may be limited largely to lowland areas. This trend is reflected in the suitability maps for the *quadramaculatus* and *marmoratus* phenotypes in 2050 and 2070, which show high suitability retained at mid and high elevations (Figure 4.8). Hence, while future predictions show a contraction of suitability in the lowlands, this represents habitat that none of the phenotypes currently occupy. Higher temperatures in the lowlands could have indirect effects on members of the *quadramaculatus* species complex if community composition changes when lowland salamanders migrate upwards to escape rising temperatures. Based on previous studies that experimentally manipulated the composition of desmognathine assemblages (see review in Bruce 2011), *quadramaculatus* may benefit

from other species moving upward because it would provide additional prey sources, while *marmoratus* and *folkertsi* could suffer from increased competition due to their smaller size.

The phenotypes may also exhibit different responses to future precipitation regimes. Higher stream salamander abundance in southern Appalachia is correlated with the positive phase of the North Atlantic Oscillation during which precipitation rates increase (Warren and Bradford 2010). Under climate projections where the effects of climate change are more potent (e.g., RCP8.5), annual precipitation is expected to either remain the same or increase over the next 50 years. Some models also predict an increase in precipitation seasonality, especially in the southern portion of the study area, including the entire range of *folkertsi*. This could lead to inconsistent stream flow and therefore negatively affect recruitment if streams periodically dry up before larvae are able to mature (Barrett et al. 2010; Lowe 2012). The aquatic larval stage in all phenotypes occurs in the same streams where adults are found, but time to metamorphosis varies by phenotype—typically two years for *folkertsi*, three years for *marmoratus*, and three to four years for *quadramaculatus* (Bruce 1985, 1988; Camp et al. 2002). However, time to metamorphosis may exhibit some degree of plasticity, with lower rainfall decreasing age and size at metamorphosis, and therefore decreasing time to maturity (Camp et al. 2000; Beachy and Bruce 2003). Hence, the shorter larval stage of *folkertsi* may provide a selective advantage under some future climate scenarios, and plasticity in maturation rates could mitigate the impacts of precipitation variability, but at the cost of smaller body size and thus increased competition from other salamanders.

While anthropogenic warming may not be an immediate threat to habitat suitability for the *quadramaculatus* species complex, anthropogenic land use changes are predicted to continue to reduce available habitat and connectivity between salamander populations, particularly around the Asheville Basin. Many populations currently reside in protected areas, but development and road traffic within these areas to accommodate tourism can negatively affect habitat quality, increase road mortality, and fragment populations (Antunes et al. 2022; Halstead et al. 2022). The protected areas in southern Appalachia are also largely disjunct, and major roads run directly through and alongside existing habitats, which can decrease habitat suitability and negatively affect dispersal (Årevall et al. 2018; Cayuela et al. 2018; Remon et al. 2022). Thus the results of this study indicate that reducing anthropogenic disturbance within protected areas, limiting the outward spread of development, and improving habitat around streams should be conservation priorities for the *quadramaculatus* complex.

Table 4.1. Number of occurrence points and background points used to build ecological niche models (ENMs). SB = background to account for spatial bias using GBIF data for *Desmognathus* (non-target species) and *Plethodon* occurrence; GR = systematic 5 km grid across combined range of all occurrence points (see text for details). Note: GR background points were not put through thinning procedure because they were already at 5 km distance.

| | Number of points | |
|-------------------|-----------------------|-----------------------|
| | After quality control | After thinning (5 km) |
| Clade | | |
| Nantahala | 158 | 95 |
| Pisgah | 182 | 102 |
| Phenotype | | |
| folkertsi | 45 | 16 |
| marmoratus | 179 | 65 |
| quadramaculatus | 850 | 284 |
| Background | | |
| Spatial bias (SB) | 1011 | 298 |
| Gridded (GR) | 3,174 | 3,174 |

Table 4.2. Bioclimatic variables used in baseline models. WC = all 19 bioclimatic variables from WorldClim v2.1; WCR = reduced correlation bioclimatic variables from WorldClim v2.1; CH = all 19 bioclimatic variables from CHELSA v1.2; CHR = reduced correlation bioclimatic variables from CHELSA.

| | Description | WC | WCR | CH | CHR |
|-------|--------------------------------------|----|-----|----|-----|
| BIO1 | Annual mean temperature | x | x | x | x |
| BIO2 | Mean diurnal range | x | x | x | |
| BIO3 | Isothermality (BIO2/BIO7) | x | | x | |
| BIO4 | Temperature seasonality | x | | x | |
| BIO5 | Max temperature of warmest month | x | | x | |
| BIO6 | Min temperature of coldest month | x | | x | |
| BIO7 | Temperature annual range (BIO5-BIO6) | x | x | x | |
| BIO8 | Mean temperature of wettest quarter | x | | x | x |
| BIO9 | Mean temperature of driest quarter | x | | x | x |
| BIO10 | Mean temperature of warmest quarter | x | | x | |
| BIO11 | Mean temperature of coldest quarter | x | | x | |
| BIO12 | Annual precipitation | x | x | x | x |
| BIO13 | Precipitation of wettest month | x | | x | |
| BIO14 | Precipitation of driest month | x | | x | |
| BIO15 | Precipitation seasonality | x | x | x | x |
| BIO16 | Precipitation of wettest quarter | x | | x | |
| BIO17 | Precipitation of driest quarter | x | | x | |
| BIO18 | Precipitation of warmest quarter | x | | x | |
| BIO19 | Precipitation of coldest quarter | x | | x | |

Table 4.3. Reclassification of land use and land cover (LULC) rasters used in ENMs and habitat stability analyses. For Sohl et al. 2018, wetlands were split into pasture and forest, but for Li et al. 2022, all wetland was reclassified as pasture.

| Sohl et al. 2018 | Li et al. 2022 | Reclassification |
|---|----------------|------------------|
| 1 Open Water | Water (8) | NA |
| 2 Urban/Developed | Urban (1) | developed |
| 3 Mechanically Disturbed National Forests | — | developed |
| 4 Mechanically Disturbed Other Public Lands | — | developed |
| 5 Mechanically Disturbed Private | — | developed |
| 6 Mining | — | developed |
| 7 Barren | Barren (9) | developed |
| 8 Deciduous Forest | Forest (4) | forest |
| 9 Evergreen Forest | Forest (4) | forest |
| 10 Mixed Forest | Forest (4) | forest |
| 11 Grassland | Grassland (6) | pasture |
| 12 Shrubland | Shrub (5) | pasture |
| 13 Cultivated Cropland | Crop (2) | cropland |
| 14 Hay/Pasture | Pasture (3) | pasture |
| 15 Herbaceous Wetland | Wetland (7) | pasture |
| 16 Woody Wetland | Wetland (7) | forest * |
| 17 Perennial Ice/Snow | — | NA |

Table 4.4. Results for “best” models for each combination of variables and backgrounds tested. Values were tested against the distribution of 1000 null models built with withheld occurrence data. Significant values indicate that the empirical model scored higher than the null (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). AUC = Average validation area under the curve (AUC); CBI = Continuous Boyce Index (CBI); Reduced = only reduced correlation bioclimatic variables included; LULC = land use and land cover variables (forest, crop, pasture, developed) and forest stand age included in models. SB = spatial bias background, GR = gridded background, F = folkertsi phenotype, M = marmoratus phenotype, Q = quadramaculatus phenotype, N = Nantahala clade, P = Pisgah clade.

| | Worldclim | | | | Worldclim reduced | | | |
|---|------------------|---------------|----------------|----------------|--------------------------|--------------|----------------|----------------|
| | SB background | | GR background | | SB background | | GR background | |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| F | 0.71 | 0.48 | 0.91* | 0.45* | 0.82 | 0.31 | 0.9* | 0.16 |
| M | 0.79 | 0.43 | 0.85*** | 0.64** | 0.73 | 0.43 | 0.96* | 0.81*** |
| Q | 0.62 | 0.71* | 0.73 | 0.82* | 0.57 | 0.48 | 0.77* | 0.86* |
| N | 0.73 | 0.41 | 0.9* | 0.68 | 0.8 | 0.39 | 0.9* | 0.61 |
| P | 0.8* | NA | 0.83* | 0.66 | 0.73 | 0.53 | 0.82* | 0.72** |
| | Worldclim + LULC | | | | Worldclim reduced + LULC | | | |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| F | 0.78 | 0.71* | 0.89* | 0.72** | 0.84 | 0.27 | 0.91* | 0.71*** |
| M | 0.71 | 0.56 | 0.85*** | 0.65** | 0.72 | 0.43 | 0.86*** | 0.68*** |
| Q | 0.57 | 0.34 | 0.73 | 0.73 | 0.62 | 0.39 | 0.77* | 0.89* |
| N | 0.79 | 0.25 | 0.9* | 0.72 | 0.80 | 0.45 | 0.91** | 0.66 |
| P | 0.72 | 0.54 | 0.84** | 0.69 | 0.72 | 0.48 | 0.82*** | 0.72* |
| | CHELSA | | | | CHELSA reduced | | | |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| F | 0.88*** | 0.55* | 0.88*** | 0.64*** | 0.71* | 0.44 | 0.73 | 0.59* |
| M | 0.7 | 0.55 | 0.82*** | 0.73*** | 0.67 | 0.53 | 0.79*** | 0.62** |
| Q | 0.66 | 0.66 | 0.76** | 0.87** | 0.66 | 0.45 | 0.71** | 0.9** |
| N | 0.77 | 0.53 | 0.85* | 0.83*** | 0.78 | 0.41 | 0.84** | 0.69** |
| P | 0.77 | 0.86 | 0.82* | 0.78* | 0.74** | 0.7** | 0.73*** | 0.57*** |
| | CHELSA + LULC | | | | CHELSA reduced + LULC | | | |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| | AUC | CBI | AUC | CBI | AUC | CBI | AUC | CBI |
| F | 0.88*** | 0.63** | 0.88*** | 0.76*** | 0.6 | NA | 0.78** | 0.51** |
| M | 0.69 | 0.41 | 0.85*** | 0.6** | 0.67 | 0.41 | 0.82*** | 0.75*** |
| Q | 0.66 | 0.53 | 0.77** | 0.87** | 0.66 | 0.63 | 0.74** | 0.9** |
| N | 0.78 | 0.46 | 0.87* | 0.71** | 0.78 | 0.61 | 0.87* | 0.77** |
| P | 0.77 | 0.86 | 0.83** | 0.79** | 0.72 | 0.62* | 0.77*** | 0.82*** |

Table 4.5. Variable permutation importance for CHELSA models with GR background, with and without LULC. Variables with <10% permutation importance in all models are excluded. Top two variables for each model in bold. Arrows indicate how variables responded to suitability according to model response curves (i.e., ↑ = increasing suitability with increase in variable; ↓ = decreasing suitability with increase in variable). Forest = forest stand age, F = folkertsi phenotype, M = marmoratus phenotype, Q = quadramaculatus phenotype, N = Nantahala clade, P = Pisgah clade.

| | Response | CHELSA + LULC | | | | | CHELSA (no LULC) | | | | |
|--------|----------|---------------|-------------|-------------|-------------|-------------|------------------|-------------|-------------|-------------|-------------|
| | | F | M | Q | N | P | F | M | Q | N | P |
| forest | ↑ | 12.6 | 33.4 | 37.7 | 17.8 | 4.7 | - | - | - | - | - |
| BIO2 | ↑ | 34.4 | 0.7 | 13.9 | 13.7 | 0.3 | 34.1 | 0 | 15.4 | 11.4 | 0 |
| BIO4 | ↓ | 0 | 0 | 18.1 | 0 | 17.6 | 0 | 0 | 11.6 | 0 | 27.5 |
| BIO7 | ↓ | 0 | 23.4 | 3.9 | 22.8 | 13.8 | 0 | 41.4 | 18.3 | 43.8 | 6.1 |
| BIO8 | ↓ | 53 | 0 | 0 | 39.5 | 9.5 | 65.9 | 0 | 0 | 33.3 | 7.2 |
| BIO11 | ↓ | 0 | 35.3 | 11 | 3.6 | 11.9 | 0 | 48.2 | 27.3 | 7.7 | 16.2 |
| BIO17 | ↑ | 0 | 0 | 0 | 0 | 19.2 | 0 | 0 | 0 | 0 | 19 |
| BIO18 | ↓ | 0 | 0 | 1.9 | 0 | 15.5 | 0 | 0 | 8.4 | 0 | 17.8 |

Table 4.6. Identity tests for best CHELSA+LULC ENMs. For *D* and *I*, 0 = no niche overlap, 1 = complete niche overlap, ** $p < 0.01$, * $p < 0.05$, ns = not significant. Significant values (based on 100 pseudoreplicates) indicate that niche space used by the two groups is not equivalent. Correlation (*R*) is the proportion of niche space shared between both groups. F = folkertsi phenotype, M = marmoratus phenotype, Q = quadramaculatus phenotype, N = Nantahala clade, P = Pisgah clade.

| | Geographic space | | | Environmental space | | |
|--------|------------------|----------|-----------|---------------------|-----------|----------|
| | <i>D</i> | <i>I</i> | <i>R</i> | <i>D</i> | <i>I</i> | <i>R</i> |
| M vs Q | 0.72** | 0.93** | 0.88 (ns) | 0.74 (ns) | 0.87 (ns) | 0.92** |
| F vs Q | 0.46** | 0.74** | 0.57 (ns) | 0.15** | 0.38** | 0.31** |
| F vs M | 0.44** | 0.7** | 0.28** | 0.17** | 0.4* | 0.31** |
| P vs N | 0.35** | 0.59** | 0.04** | 0.29** | 0.44** | 0.65** |

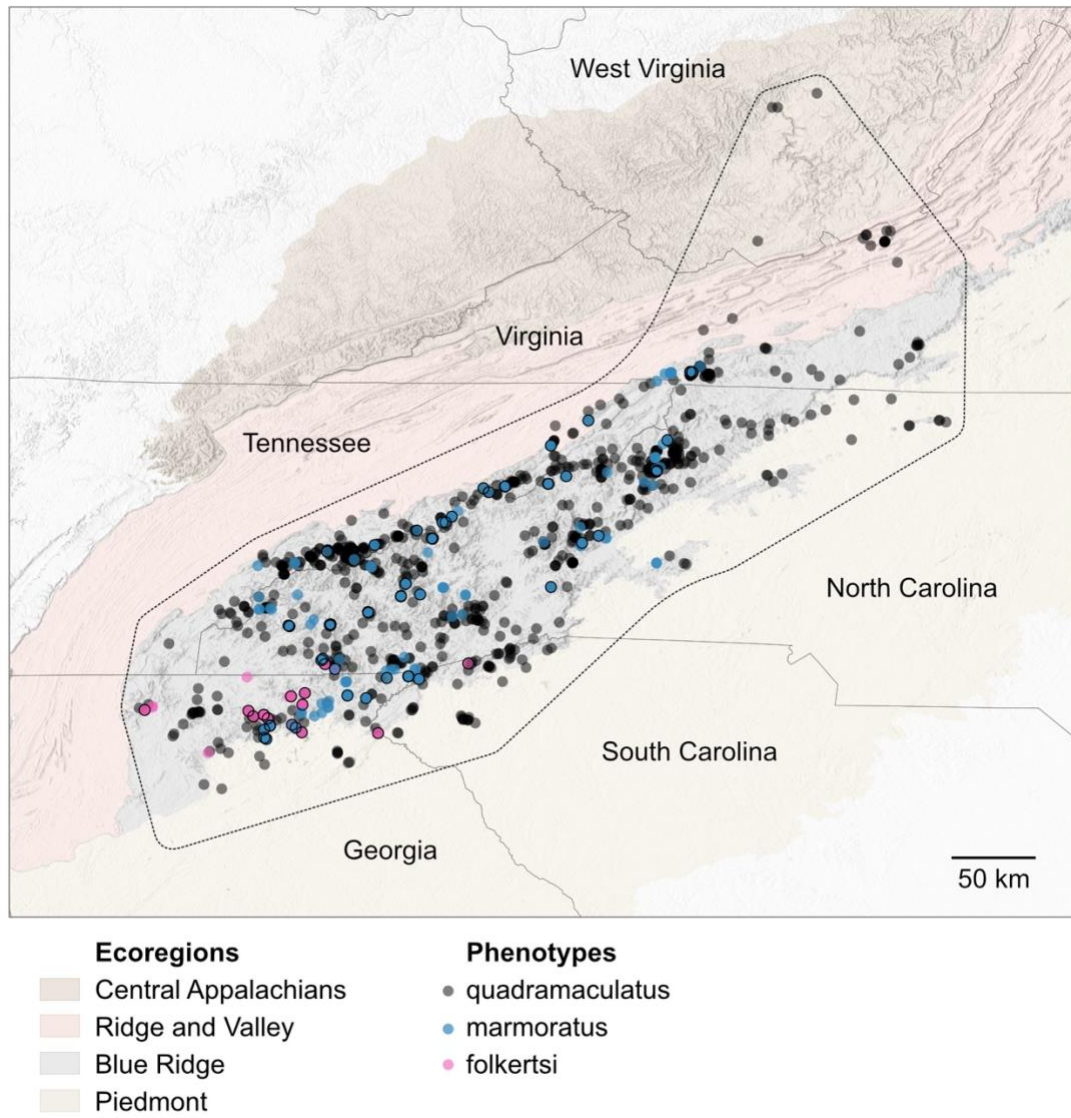


Figure 4.1. Shaded relief map of study area and occurrence locations where individuals of the *Desmognathus quadramaculatus* species complex have been found within the last 30 years (1992-2022). Black dotted line indicates boundary of study area used for analyses. Occurrences are colored by phenotype; sites where two phenotypes were reported in the same location are outlined in black. The four ecoregions where the species occur are highlighted. Solid grey lines indicate state boundaries.

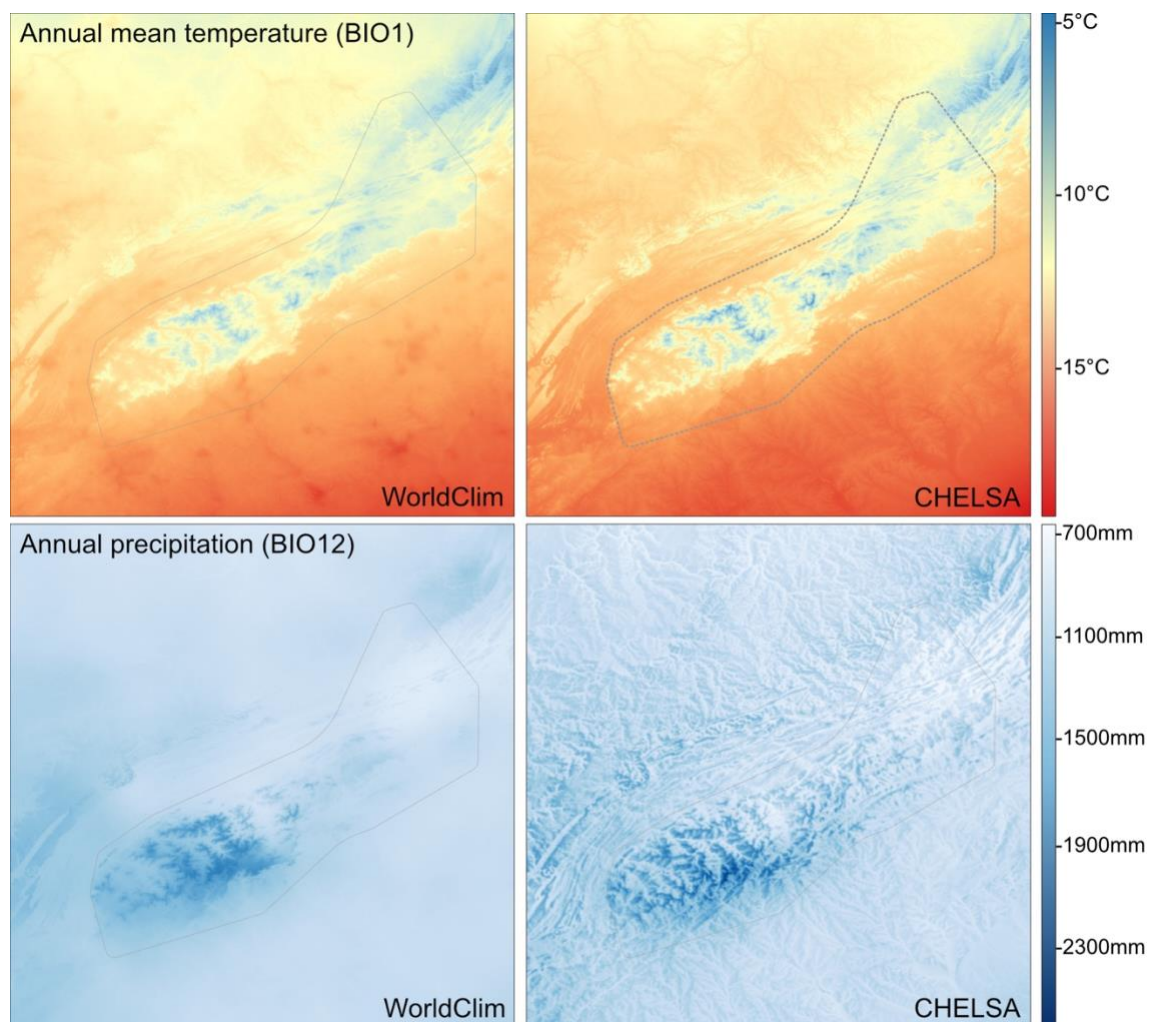


Figure 4.2. Comparison of two bioclimatic variables for Worldclim 2.1 and CHELSA 1.4. BIO1 (annual temperature, top); BIO12 (annual precipitation, bottom).

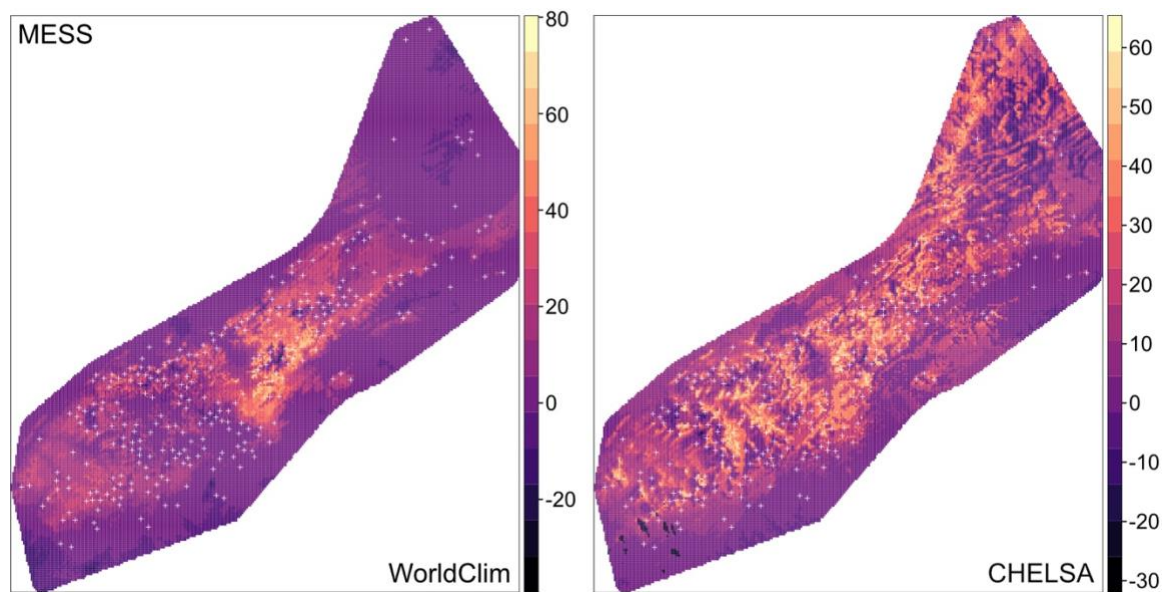


Figure 4.3. Multivariate Environmental Similarity Surfaces (MESS). WorldClim v2.1 (left) and CHELSA v1.2 (right) for all bioclimatic variables. White crosses indicate occurrence points. Areas with high positive values are more similar to climate conditions found at occurrence points, while areas with high negative values are more dissimilar.

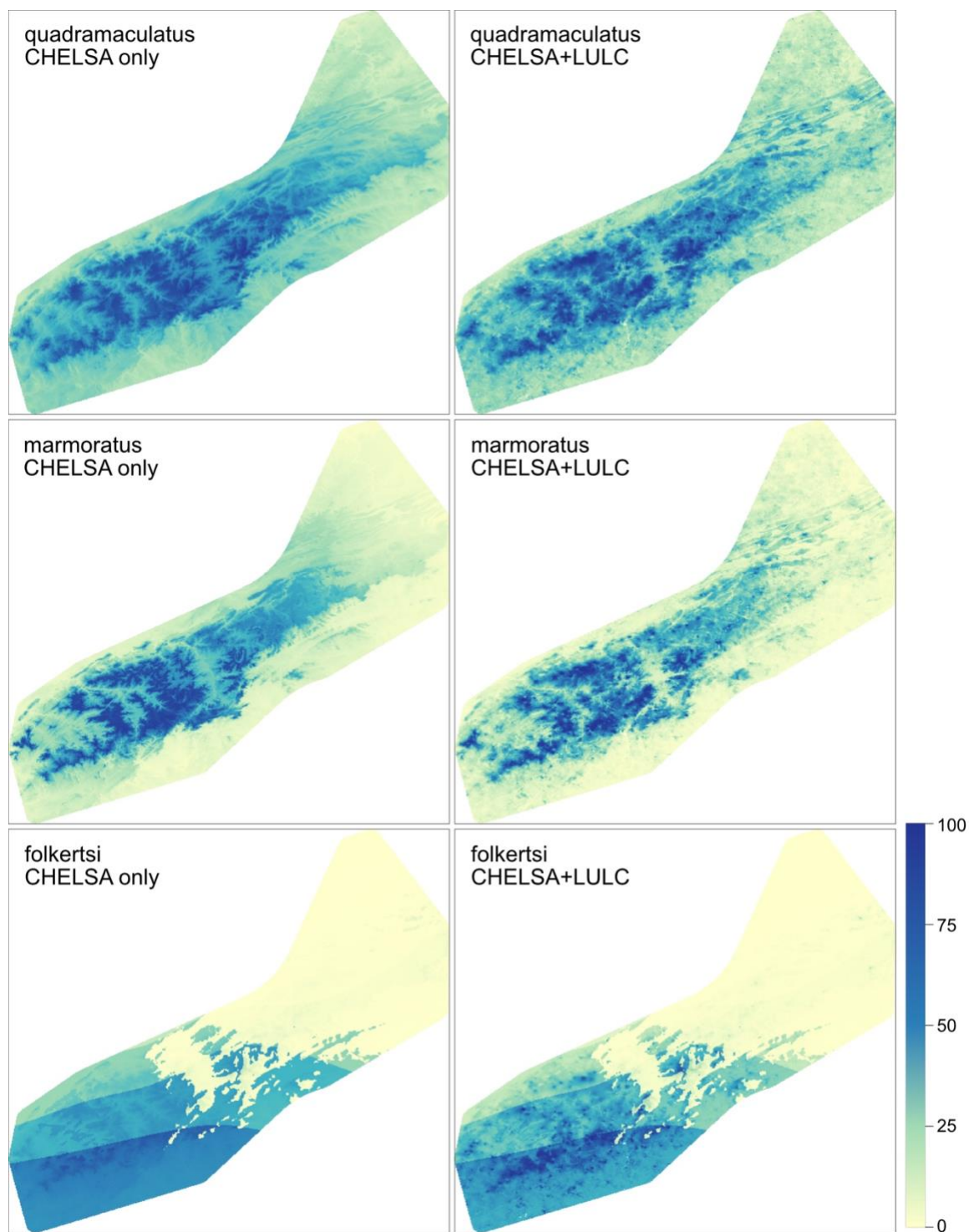


Figure 4.4. Habitat suitability maps for phenotypes. *Quadramaculatus* (top), *marmoratus* (middle), and *folkertsi* (bottom) for best CHELSA models with bioclimatic variables only (left) and with bioclimatic variables and land use and land cover (LULC, right). Darker colors indicate higher percent habitat suitability.

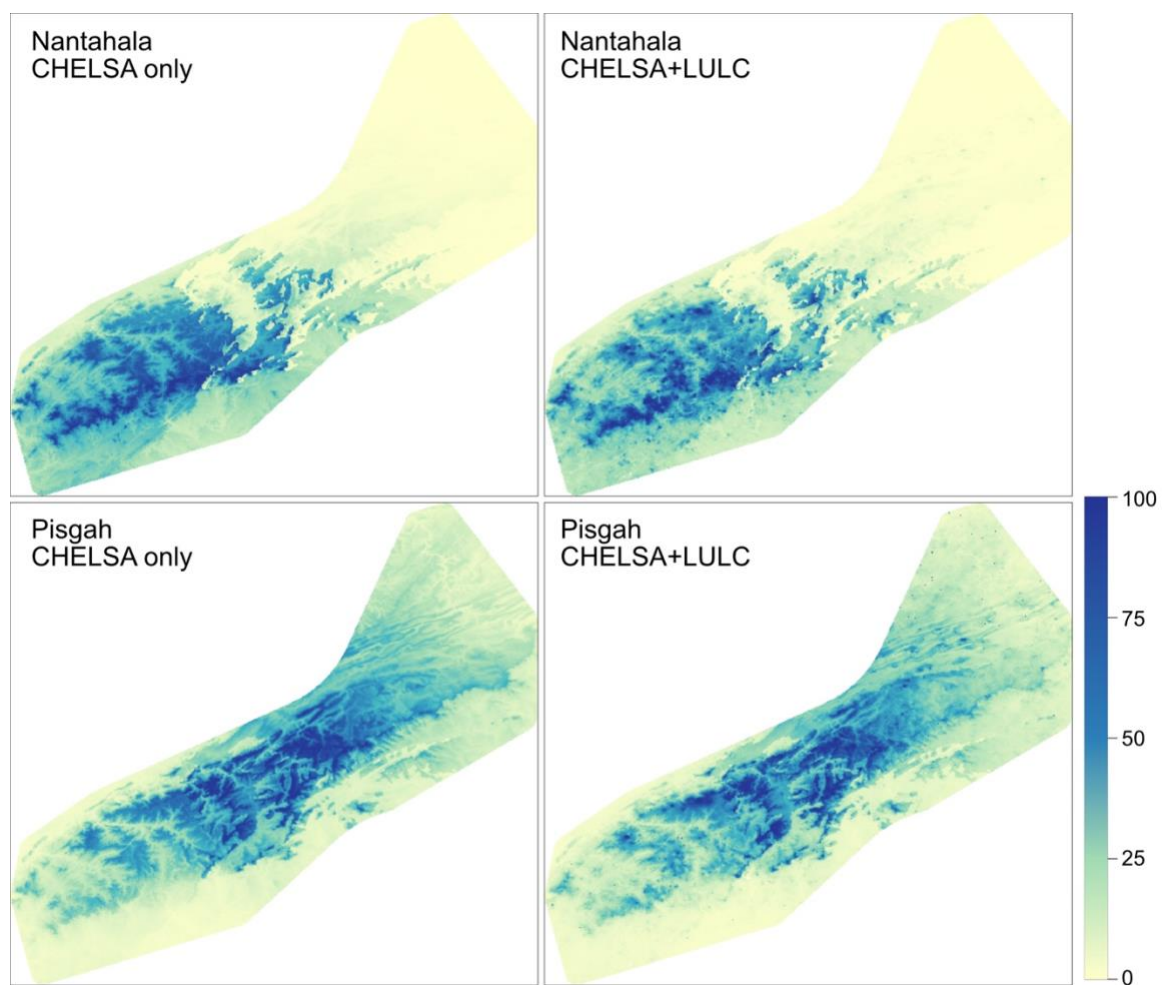


Figure 4.5. Habitat suitability maps for genetic clades. Nantahala (top) and Pisgah (bottom). Variables and colors follow Figure 4.4.

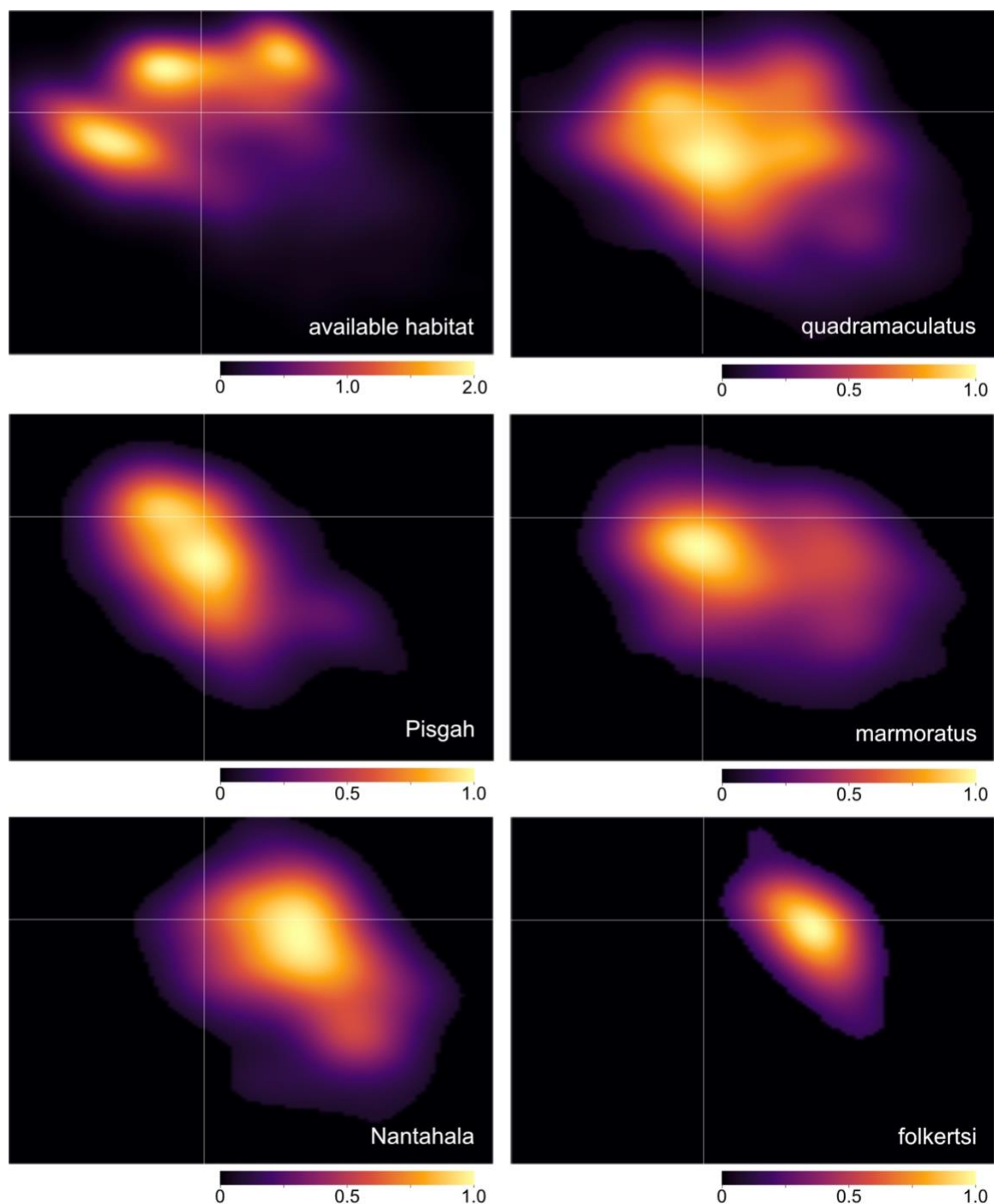


Figure 4.6. Visualization of results for identity by environment tests for CHELSA+LULC. Axes represent correlated variables. Top left shows all available habitat, while the rest show occurrence in environment space for each clade or phenotype. White lines indicate zero on both axes; axis dimensions are the same for all graphs. Colors denote density. Note that color scale for available habitat differs from the other graphs. All comparisons between clades or phenotypes were significantly different except for *marmoratus* versus *quadramaculatus*. See Table 4.6 for full results.

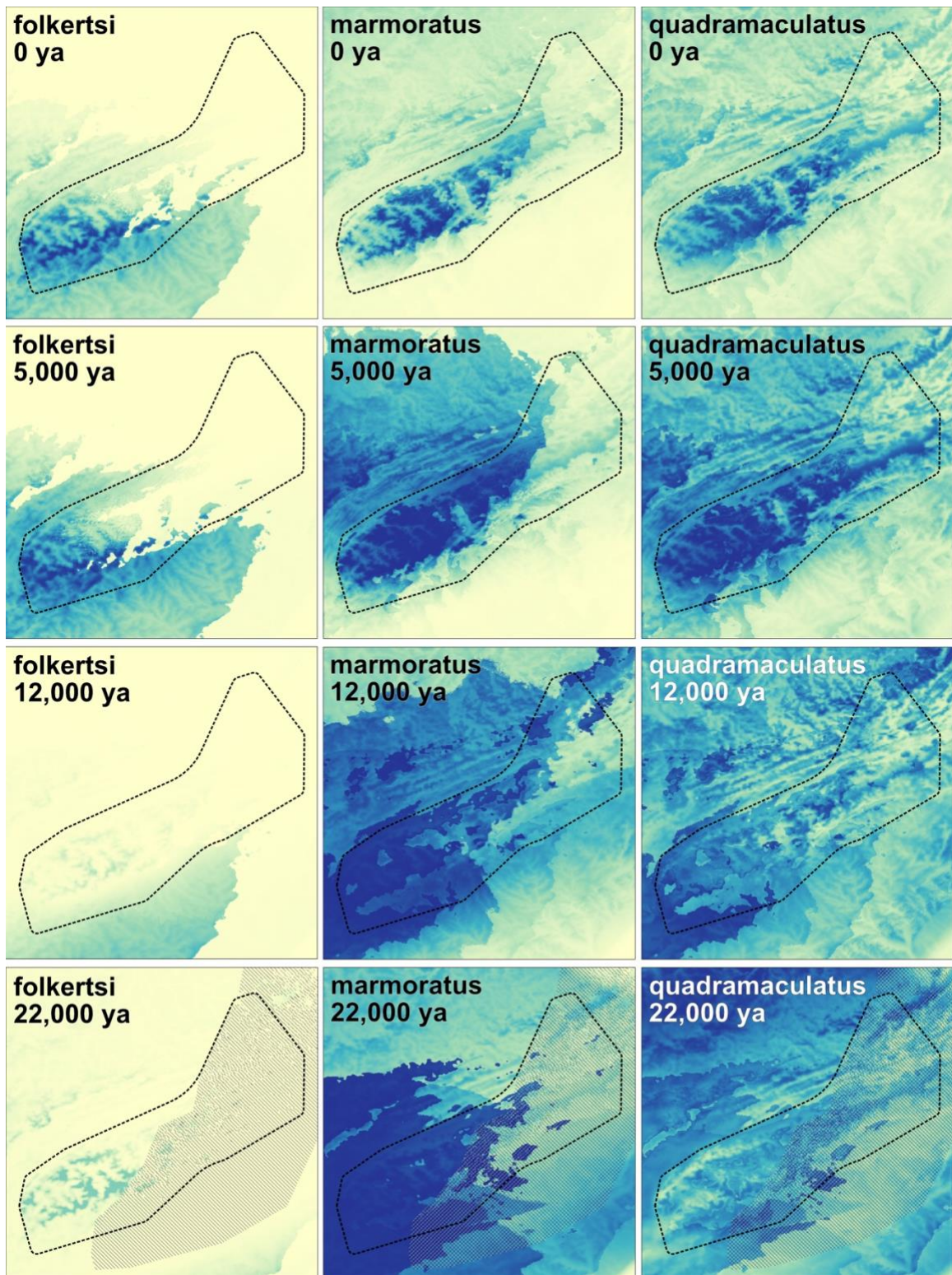


Figure 4.7. Habitat suitability maps for phenotypes projected to the Last Glacial Maximum (LGM). *Folkertsi* (left), *marmoratus* (middle), and *quadramaculatus* (right) for present-day (0 ya, top), 5 kya, 12 kya, and LGM (~22 kya, bottom). Hashed lines on 22 kya show projected extent of discontinuous permafrost during LGM. Darker colors indicate higher percent habitat suitability. (Continued on next page.)

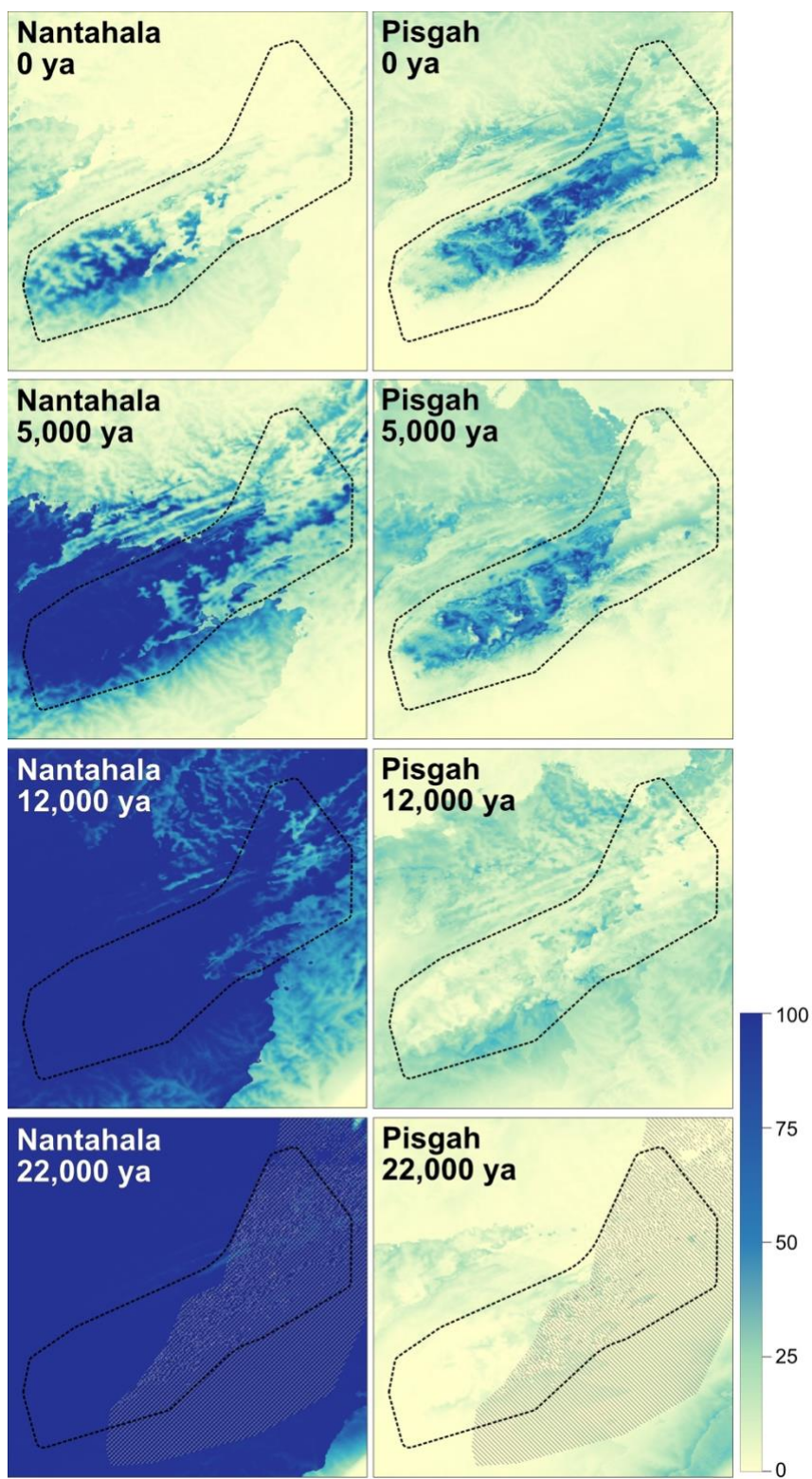


Figure 7 (continued)

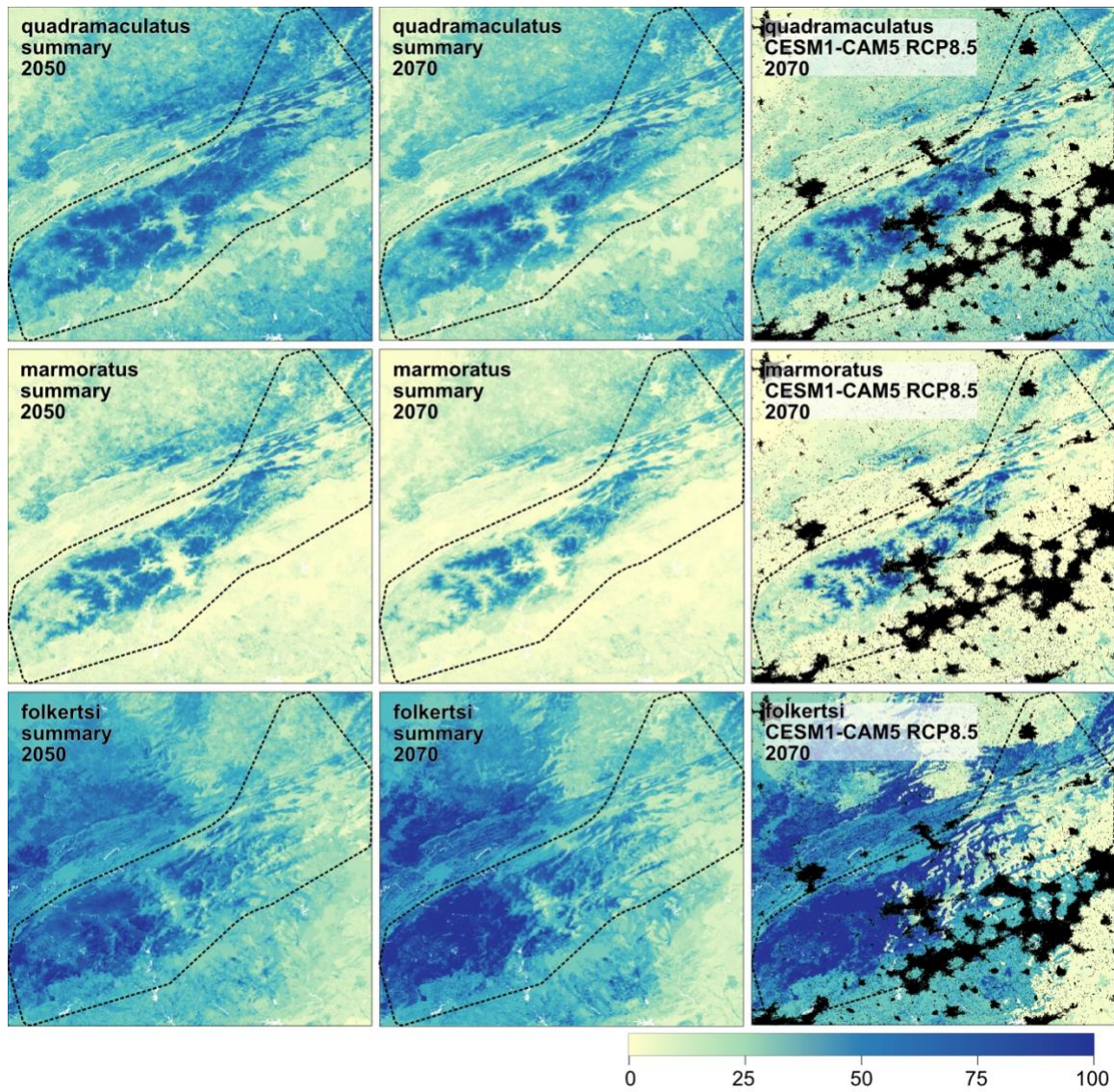


Figure 4.8. Habitat suitability maps for phenotypes projected to 2050 and 2070. Weighted summaries of habitat suitability for phenotypes using all six CHELSA+LULC climate projections for 2050 (left) and 2070 (center). Projection under CESM1-CAM5 GCM and RCP8.5 emissions scenario for 2070, overlaid with urban areas in black (right). This scenario represents the “worst case” (i.e., largest area of lowest suitability) of the predictions. See text for model details.

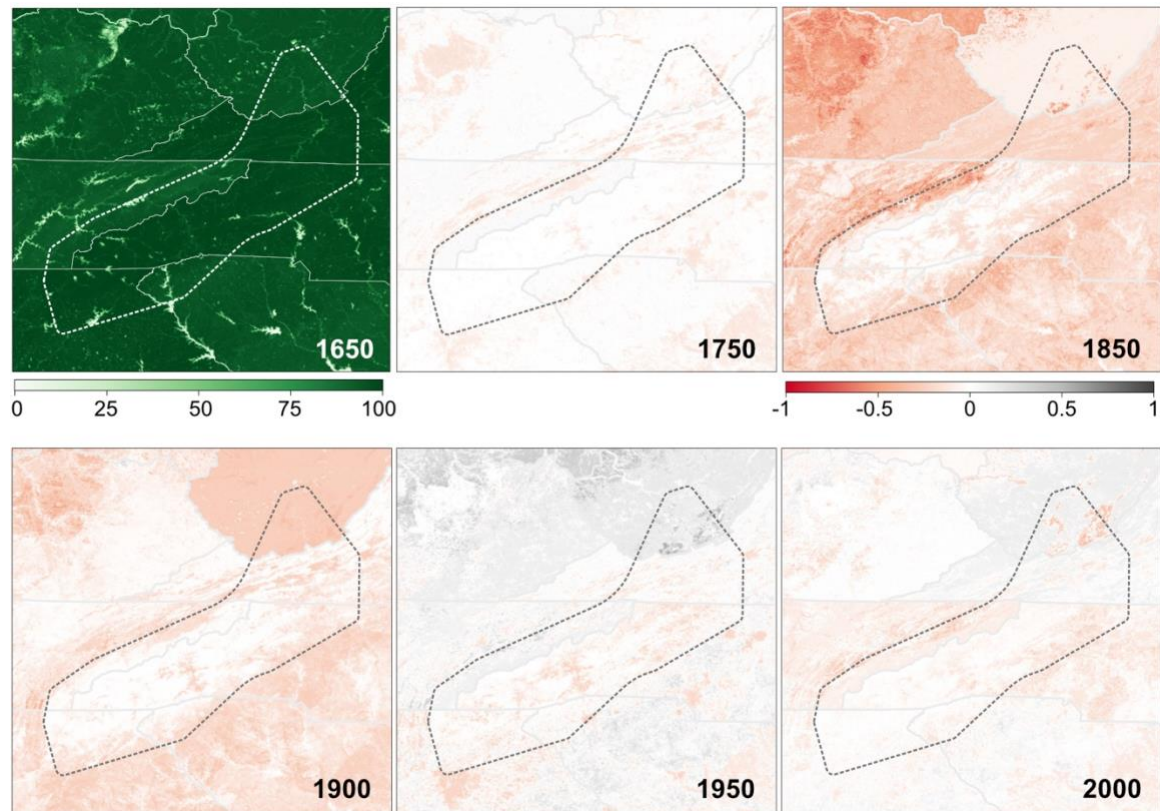


Figure 4.9. Percent forest cover in study region from 1650-2000. Six time points shown: 1650 shows percent forest cover, while each subsequent time point shows percent forest change from previous time period (e.g., 1750 shows difference between 1750 and 1650, and 1850 shows difference between 1850 and 1750) on a scale from -1 (100% forest lost) to 1 (100% forest gained).

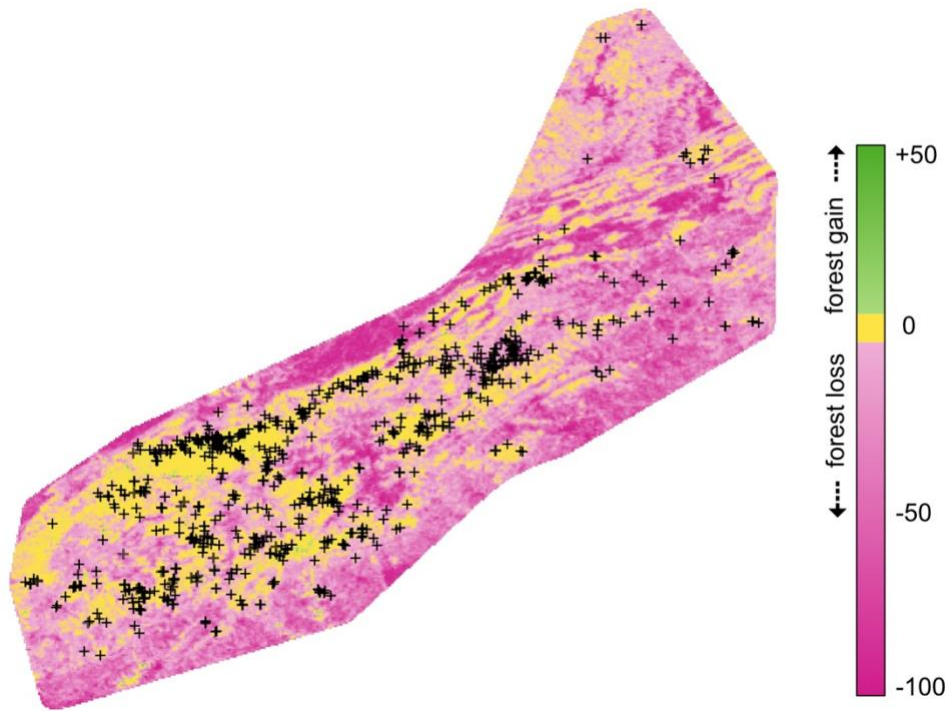


Figure 4.10. Difference in percent forest cover between 1650 and 2000. Yellow areas have the same net forest cover in 2000 as in 1650, green areas have a higher percent cover than 1650, and pink areas lower percent forest cover than 1650.

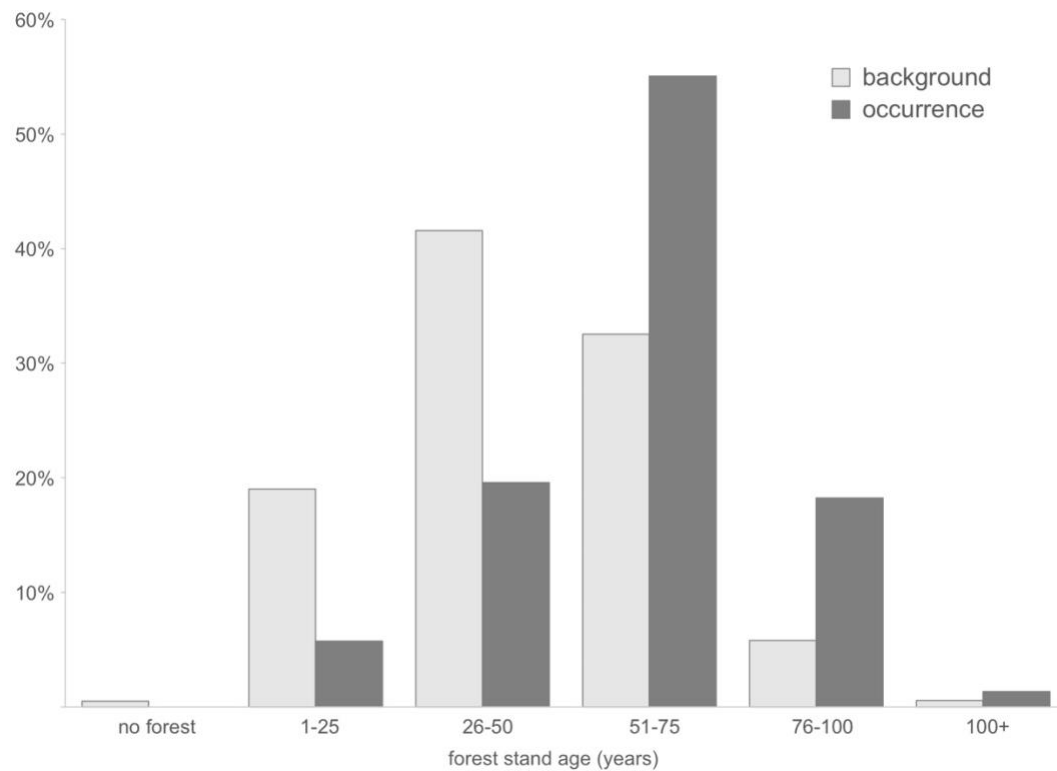


Figure 4.11. Forest stand age found at 5 km grid background points and occurrence points.

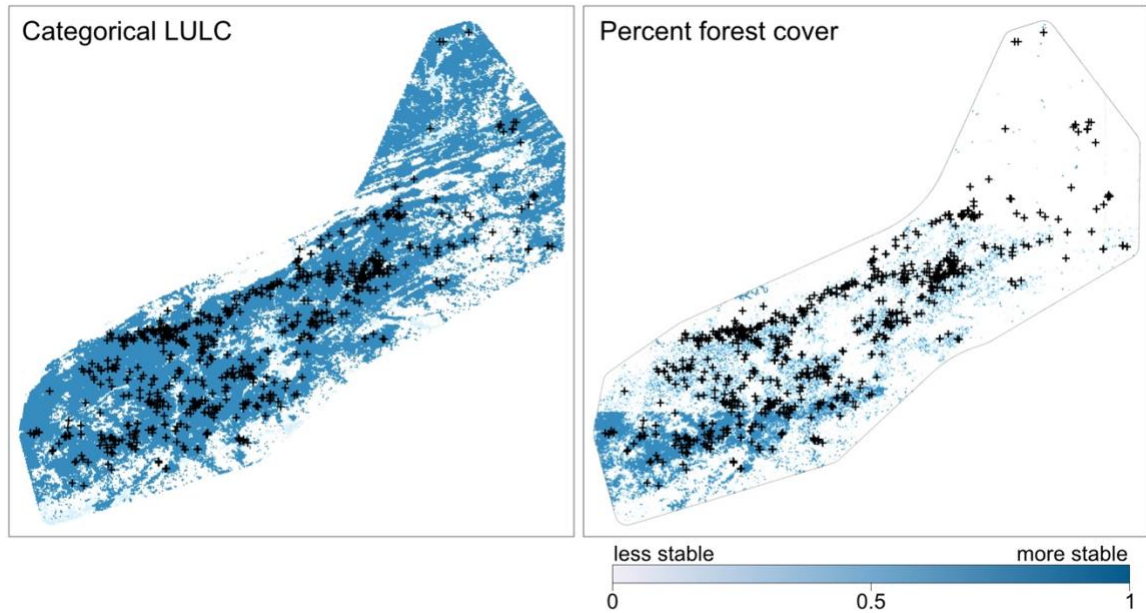


Figure 4.12. Habitat stability from 1650-2000. Habitat stability measured as binary change for categorical land use and land cover (LULC: forest, crop, pasture, or developed; left) and change in percent forest cover (0-25%, >25-50%, >50-75%, >75%; right) between 1650 and 2000. Black crosses are occurrence points. High numbers (close to 1) indicate that the land cover has stayed the same over most of the time span, while low numbers (close to 0) indicate that land cover has switched between categories several times (i.e., forest cover to agriculture, or >75% forest cover to >50-75% forest cover) over the time span.

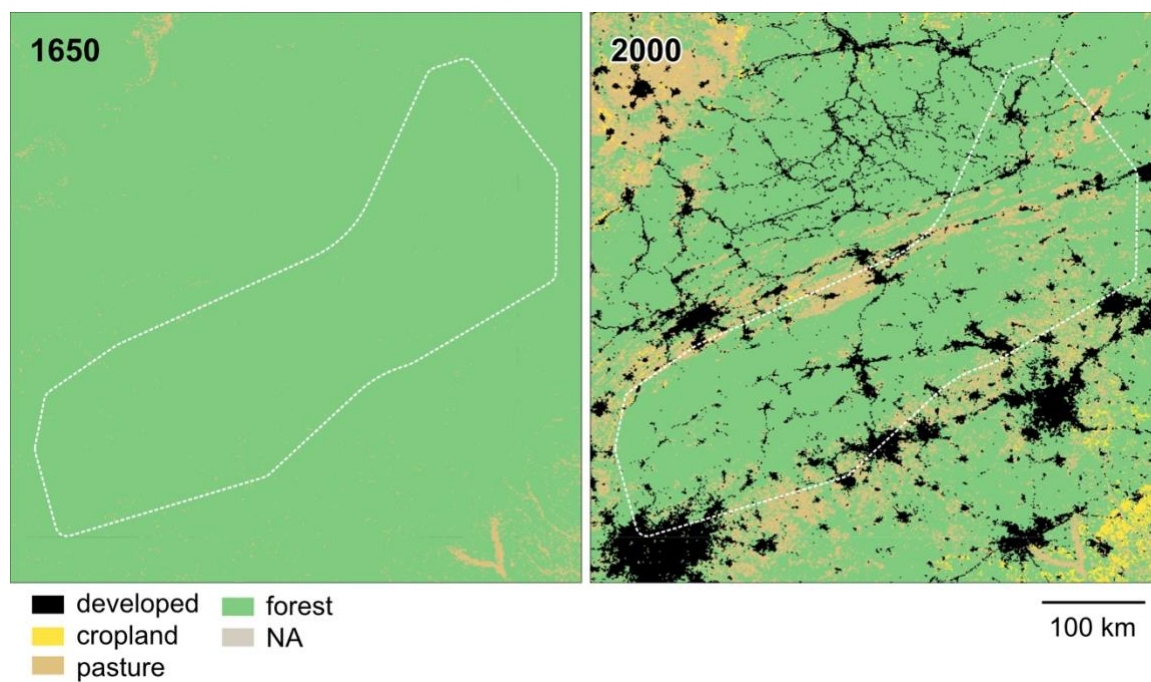


Figure 4.13. Land use and land cover for 1650 and 2000. Study region outlined with white dashed line.

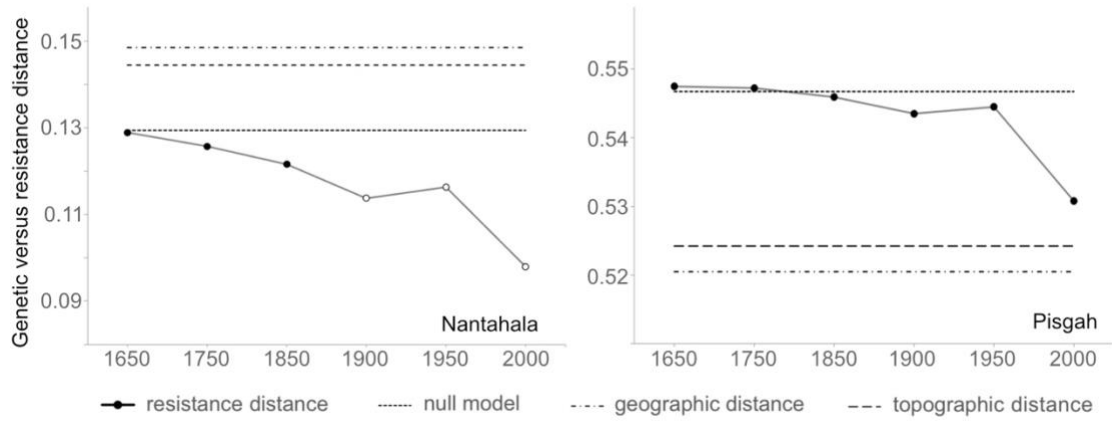


Figure 4.14. Correlation between genetic distance and resistance distance between 1650 and 2000. Six time points shown for Nantahala (left) and Pisgah (right). Closed circles indicate significant correlation, while open circles are not significant (see text for details). Null model (100% forest cover in all pixels), geographic distance, and topographic distance included for comparison. Note difference in scale between y-axes.

CHAPTER 5. CONCLUSION

In my dissertation, I provided evidence that members of the *Desmognathus quadramaculatus* species complex are divided into two independently evolving cryptic lineages, which I designated Nantahala and Pisgah after the National Forests where they are found. These lineages are supported by a multitude of diagnostic alleles that can be used to genetically differentiate between them, along with supporting evidence from phylogenetic and population structure analyses that show that there is no admixture at sites where the lineages overlap geographically. Finally, though there is geographic overlap between the lineages, they appear to occupy different ecological niches, have different evolutionary histories, and are projected to respond differently to future climate change. Hence these two lineages could be considered separate species under the Phylogenetic, Evolutionary, Ecological, and Biological Species Concepts.

The nominal species *D. quadramaculatus* and *D. marmoratus*, as previously delimited by morphological differences, need to be revised to align the taxonomy with what we now know about the evolutionary history of this group. However, there are still many questions that remain about proposed candidate species within these lineages, and any changes should be implemented with caution so that the updated taxonomy is accurate. Below is a summary of the findings within each lineage focusing on implications for recent taxonomic changes and suggestions for further areas of study.

5.1 Evidence for divergence within the Nantahala lineage

Each of the candidate species in Nantahala form reciprocally monophyletic clades with relatively high branch support, have >50 diagnostic sites, and are supported by some of the population genetic analyses. The candidate species quadramaculatus A and F were recently described as new species by Pyron and Beamer (2022) as *D. ampileucus* and *D. gvnigeusgwotli*, respectively. *D. folkertsi* was previously described as a species by Camp et al. (2002), while marmoratus B is currently undescribed. The results of this dissertation provide support for each of these species under the Phylogenetic and Evolutionary Species Concepts, though further evidence would be needed to determine whether the described species adheres to the Biological Species Concept (see below).

However, these results raise several questions about historical and ongoing gene flow within this group. First, two individuals do not cluster with their clade in the COI haplonet analysis, where one member of marmoratus B groups with quadramaculatus A, which likely represents an introgression event. Furthermore, *D. folkertsi* and marmoratus B do not form a separate cluster at low values of K in the sNMF analysis, nor do they cluster into separate groups in the PCA. They also have fuzzy boundaries in the coMA analysis. While Pyron et al. (2022) tested whether shared alleles were due to introgression or ILS using f_4 analyses, there were not enough independent markers used in those analyses to statistically distinguish between the two processes. Replicating those analyses with larger sample sizes and sufficient markers would be useful for determining whether there is recent or ongoing gene flow among groups in Nantahala.

Second, the geographic boundary for quadramaculatus F is unclear, likely due to insufficient sampling. Jackson (2005) identified a mtDNA clade that is equivalent to

quadramaculatus F, but it extends farther than samples from other studies, leading to more geographic overlap with quadramaculatus A. It is uncertain whether this mitonuclear discordance is due to limited sampling in nuclear studies that did not catch more individuals in quadramaculatus F, or whether the quadramaculatus F mitochondrial haplotypes extend beyond the borders of the nuclear haplotype. This pattern may indicate that quadramaculatus F recently underwent a demographic expansion and is displacing quadramaculatus A south of the Great Smoky Mountains, or that there is a narrow hybrid zone along the border of the two groups that has led to asymmetrical mitochondrial introgression from quadramaculatus F to A (Toews and Brelsford 2012). Improved sampling in and around the Great Smoky Mountains, along with demographic analyses testing for signatures of population expansion and tests of introgression would help clarify species boundaries and provide further insight into the evolutionary history of these taxa.

Lastly, there is disagreement as to the status of the marmoratus phenotype outside of the range of marmoratus B. Pyron et al. (2020) claimed that I misidentified individuals in the Nantahala clade as having the marmoratus phenotype in Chapter 2 (Jones and Weisrock [2018]), and that all marmoratus phenotypes in that region are members of the Pisgah clade. According to GBIF data (Chapter 4), marmoratus phenotypes can be found in many areas around Nantahala that have not currently been sampled in any genetic studies using nuclear data. Hence improved sampling of the marmoratus phenotype in Nantahala would be useful for determining where the southernmost boundary of the Pisgah clade lies, how far the range of marmoratus B extends, and whether there is any overlap in occurrence between marmoratus phenotypes from both clades.

5.2 Evidence for divergence within the Pisgah lineage

There is considerable disagreement among studies as to the status of candidate species within Pisgah, especially regarding whether phenotype divergence is congruent with genetic divergence. For example, Pyron et al. (2020) compared the same individuals using phylogenies inferred from a mitochondrial marker and 381 AHE nuclear loci. In the mitochondrial phylogeny, quadramaculatus and marmoratus individuals were interdigitated within two reciprocally monophyletic clades (quad/marmDE and quad/marmCGH). In the nuclear phylogenies, the same individuals were split slightly differently between three reciprocally monophyletic clades, with one individual (8408, designated “marmoratus H”) grouped in a non-reciprocally monophyletic clade composed only of marmoratus E. In Chapter 3, I found that while individuals were grouped by phenotype in the phylogeny, the clades were not supported, and mitochondrial haplotype networks and nuclear haplowebs both showed interdigitation of phenotypes. Additionally, CoMa analyses indicated that all the individuals in Pisgah comprised one field for recombination, suggesting that the proposed candidate species are not reproductively isolated. The extent of mitonuclear discordance indicates either that there is currently (or was previously) extensive introgression between quadramaculatus and marmoratus or that the nuclear tree poorly reflects evolutionary relationships. While Pyron et al. (2022) suggested the former, the weight of total evidence suggests the latter.

Furthermore, the proposed taxonomy in Pyron and Beamer (2022), which separates the quadramaculatus phenotypes in Pisgah into two species (quadramaculatus D as *D. kanawha*, and quadramaculatus ECG as *D. mavrokoilius*, respectively) while excluding marmoratus, is largely unsupported by the currently available data. I found

limited support for quadramaculatus D as genetically differentiated from the rest of the Pisgah clade, and ENMs from Chapter 4 indicated that the far northern portion of the range where this group is located likely contained little to no suitable habitat until recently. This indicates that quadramaculatus D may represent a population that formed as members of the Pisgah clade expanded post-LGM, with the few fixed alleles in this group possibly arising as a consequence of genetic drift (i.e., founder effects). Considering the low branch support in the nuclear phylogenies, and clade switching between phylogenies, this group does not qualify as a species even under the relatively generous PSC.

The data also does not support delimiting quadramaculatus ECG as a species, especially considering that the grouping excludes marmoratus C, with which it has extensive sharing of alleles. Furthermore, none of the genetic markers that Pyron and Beamer (2022) use to delimit the species are diagnostic because they excluded members of quadramaculatus C and G during the creation of COI barcodes. They also claim that there are fixed differences in the nuclear loci used in Pyron et al. (2022), but did not publish them. I did not find any fixed markers for this group despite much of the data between our studies overlapping.

Other hypotheses for genetic structuring in Pisgah emerge when considering all the available data. For example, Beamer and Lamb (2020) proposed three candidate species from clades based on a COI phylogeny: (1) quadramaculatus D, (2) quadramaculatus/marmoratus C, and (3) quadramaculatus/marmoratus E. The BEAST phylogeny using ddRAD data from Chapter 2 (which does not contain mitochondrial loci) supports the latter two clades. (It is uncertain whether quadramaculatus D would

also be supported with ddRAD data since that region was not sampled in the ddRAD data set.) It was also difficult to directly map pre-existing mitochondrial data from Jackson (2005) directly to the other data sets for Pisgah, so it is unclear if the CYTB data from that study support the COI/ddRAD clades, the AHE clades, or neither.

Data type alone can infer different phylogenetic relationships (Reddy et al. 2017), though it is unclear how much this contributed to the discrepancies between the results in the COI, AHE, and ddRAD phylogenies, particularly since there is disagreement between several of the AHE trees and some overlap between the ddRAD and AHE data sets. Furthermore, while the AHE data has denser taxon sampling, the ddRAD data set included sites that were not sampled in the AHE studies, and unsampled populations can produce differences in phylogenies and population structure analyses. Thus, future studies should specifically test whether data type (i.e., exon, intron, etc.), particular loci, or differences in taxon sampling are disproportionately affecting analyses. However, considering the short branch lengths in Pisgah and likely recent divergence time within the group, it may be impossible to resolve the history of this group using phylogenetic inference (Hahn and Nakhleh 2016).

Finally, phenotype is not a good predictor of genetic divergence within the *Desmognathus quadramaculatus* complex and should not be used as the basis of species delimitation in this group. There is definitive evidence of cryptic speciation, with the Pisgah and Nantahala clades forming species-level divergence but housing two shared morphotypes. While there are strong associations between phenotype and some geographical areas, there is also evidence that members of some phenotypes are genetically undifferentiated. While morphology can undoubtedly be connected with

ecological differentiation in some taxa, this remains a hypothesis for which there is little supporting evidence for this species complex. Fine-scale morphological differences are also likely indistinguishable to salamanders themselves, so it makes little sense to rely on them as potential markers of reproductive isolation and species divergence. The logical next step would be to focus specifically on identifying whether there are genetic markers associated with particular phenotypes, though this goal will likely remain out of reach until we have an assembled genome for *Desmognathus*.

APPENDICES

APPENDIX A. SUPPLEMENTARY INFORMATION FROM JONES AND WEISROCK (2018)

Table A1. Number of *D. quadramaculatus*, *D. marmoratus*, and *D. carolinensis* at each sample site.

| Site ID | Latitude | Longitude | <i>D. quadramaculatus</i> | <i>D. marmoratus</i> | <i>D. carolinensis</i> |
|---------|----------|-----------|---------------------------|----------------------|------------------------|
| 1 | 36.281 | -81.721 | — | 1 | — |
| 2 | 35.924 | -82.780 | 1 | — | 1 |
| 3 | 35.907 | -82.790 | 2 | — | — |
| 4 | 35.803 | -82.353 | 3 | — | 1 |
| 5 | 35.709 | -82.394 | 2 | — | — |
| 6 | 35.750 | -82.225 | 3 | — | — |
| 7 | 35.715 | -82.190 | 4 | 2 | — |
| 8 | 35.682 | -82.201 | 2 | 1 | 1 |
| 9 | 35.783 | -83.102 | 2 | — | — |
| 10 | 35.761 | -82.991 | — | 1 | — |
| 11 | 35.769 | -82.971 | 1 | 1 | — |
| 12 | 35.372 | -82.782 | 1 | — | — |
| 13 | 35.331 | -82.791 | 1 | 1 | — |
| 14 | 35.293 | -82.800 | 2 | — | — |
| 15 | 35.249 | -82.865 | — | 1 | — |
| 16 | 35.593 | -83.387 | 1 | 1 | — |
| 17 | 35.464 | -83.525 | 1 | — | — |
| 18 | 35.336 | -83.377 | 2 | 2 | — |
| 19 | 35.361 | -83.926 | 1 | 2 | — |
| 20 | 35.265 | -83.583 | 1 | — | — |
| 21 | 35.080 | -83.143 | 2 | — | — |
| 22 | 35.025 | -83.239 | 1 | — | — |
| 23 | 35.042 | -83.554 | 2 | — | — |
| 24 | 35.069 | -83.631 | 3 | — | — |
| 25 | 34.793 | -83.912 | 1 | — | — |
| 26 | 34.708 | -83.913 | — | 1 | — |
| 27 | 34.824 | -83.733 | 2 | — | — |
| 28 | 34.808 | -83.121 | 1 | — | — |
| 29 | 35.160 | -84.469 | 1 | — | — |

Table A2. Number of loci removed by major filtering parameters in iPyrad for the ingroup and outgroup data sets.

| Filtering parameter | Ingroup data set | | Outgroup data set | |
|---------------------------------------|------------------|------------------|-------------------|------------------|
| | # loci removed | # loci remaining | # loci removed | # loci remaining |
| Total loci before filtering | | 246129 | | 261148 |
| Duplicate loci | 5724 | 240405 | 5852 | 255296 |
| Maximum indels | 15904 | 224501 | 16871 | 238425 |
| Maximum SNPs per locus | 8725 | 215776 | 10265 | 228160 |
| Maximum heterozygous sites (paralogs) | 60 | 215716 | 33 | 228127 |
| Minimum number of individuals | 209068 | 6648 | 221746 | 6381 |
| Maximum number of alleles | 3343 | 3305 | 3175 | 3206 |
| Total loci post-filtering | | 3305 | | 3206 |

Table A3. Parameters for the best models for each background size from ENMeval. Best models determined by the lowest AICc for all models tested within each background class. The best overall model for both Nantahala and Pisgah had a background size of 30 km. Feature classes are mathematical transformations of climate variables: L = linear, Q = quadratic, LQ = linear + quadratic. The regularization value penalizes overparameterization to prevent overfitting, with higher values representing greater penalization. The number of parameters is the total number of parameters included in the model; features that equal zero are not included. AUC is the Receiving Operating Characteristic Under the Curve and roughly represents the accuracy of a model, with higher AUC values representing better fit. The mean AUC difference is a measure of how much “better” the AUC for the given model compares to the mean AUC across all models within the same background class.

| Background size (km) | AICc | Feature classes | Regularization value | # of parameters | AUC | Mean AUC difference |
|----------------------|---------------|-----------------|----------------------|-----------------|-------------|---------------------|
| Nantahala | | | | | | |
| 5 | 5015.2 | L | 2 | 7 | 0.58 | 0.06 |
| 10 | 4785.0 | L | 2.5 | 8 | 0.65 | 0.03 |
| 15 | 4742.0 | L | 3.5 | 7 | 0.69 | 0.02 |
| 20 | 4717.1 | L | 3.5 | 6 | 0.73 | 0.02 |
| 25 | 4589.1 | L | 2.5 | 8 | 0.78 | 0.02 |
| 30 | 4560.6 | L | 1.5 | 12 | 0.81 | 0.02 |
| 35 | 4562.5 | LQ | 2.5 | 12 | 0.82 | 0.02 |
| 40 | 4568.7 | LQ | 2.5 | 11 | 0.82 | 0.02 |
| Pisgah | | | | | | |
| 5 | 2571.6 | L | 10 | 4 | 0.62 | 0.04 |
| 10 | 2501.7 | LQ | 10 | 6 | 0.69 | 0.04 |
| 15 | 2285.3 | LQ | 4 | 12 | 0.73 | 0.04 |
| 20 | 2267.8 | LQ | 5 | 13 | 0.74 | 0.02 |
| 25 | 2141.8 | LQ | 5.5 | 10 | 0.78 | 0.03 |
| 30 | 2100.6 | LQ | 4 | 12 | 0.82 | 0.04 |
| 35 | 2189.0 | LQ | 10 | 9 | 0.81 | 0.04 |
| 40 | 2107.8 | LQ | 5.5 | 13 | 0.85 | 0.02 |

Table A4. Results of ecological niche models for Nantahala and Pisgah. Top three values contributing the highest percentage to each model, and the top three values with the most importance, are in bold.

| Climate variable | | Nantahala | | Pisgah | |
|------------------|----------------------------------|----------------------|------------------------|----------------------|------------------------|
| | | Percent contribution | Permutation importance | Percent contribution | Permutation importance |
| BIO1 | Annual Mean Temperature | 0.0 | 0.0 | 0.0 | 0.0 |
| BIO2 | Mean Diurnal Range | 0.1 | 3.2 | 7.7 | 0.0 |
| BIO3 | Isothermality | 12.5 | 12.5 | 3.1 | 0.2 |
| BIO4 | Temperature Seasonality | 4.5 | 5.9 | 8.6 | 14.4 |
| BIO5 | Max Temp. of Warmest Month | 5.3 | 0.0 | 29.1 | 0.0 |
| BIO6 | Min Temp. of Coldest Month | 12.4 | 1.2 | 0.9 | 4.6 |
| BIO7 | Temperature Annual Range | 0.2 | 0.0 | 4.8 | 0.0 |
| BIO8 | Mean Temp. of Wettest Quarter | 5.9 | 2.1 | 0.3 | 2.9 |
| BIO9 | Mean Temp. of Driest Quarter | 0.1 | 0.8 | 13.0 | 1.0 |
| BIO10 | Mean Temp. of Warmest Quarter | 9.4 | 0.0 | 9.7 | 0.0 |
| BIO11 | Mean Temp. of Coldest Quarter | 0.0 | 0.0 | 0.2 | 4.7 |
| BIO12 | Annual Precipitation | 3.3 | 0.0 | 0.1 | 0.0 |
| BIO13 | Precipitation of Wettest Month | 0.0 | 0.0 | 1.4 | 11.6 |
| BIO14 | Precipitation of Driest Month | 2.1 | 10.5 | 0.0 | 0.0 |
| BIO15 | Precipitation Seasonality | 0.0 | 0.0 | 5.7 | 1.4 |
| BIO16 | Precipitation of Wettest Quarter | 2.5 | 0.0 | 0.0 | 0.0 |
| BIO17 | Precipitation of Driest Quarter | 5.9 | 0.0 | 1.0 | 0.0 |
| BIO18 | Precipitation of Warmest Quarter | 33.0 | 63.9 | 9.1 | 48.5 |
| BIO19 | Precipitation of Coldest Quarter | 2.9 | 0.0 | 5.3 | 10.8 |

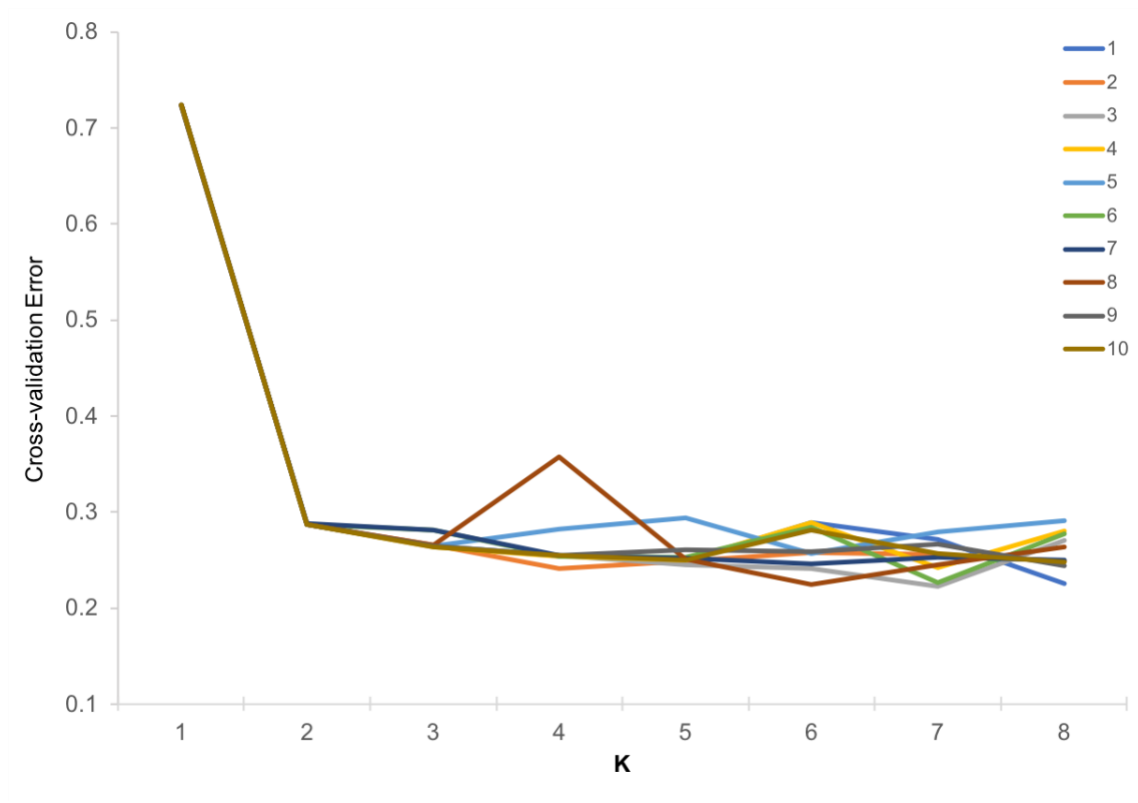


Figure A1. Graph of cross-validation error for 10 runs of Admixture for $K = 1-8$ showing an optimal K of 2.

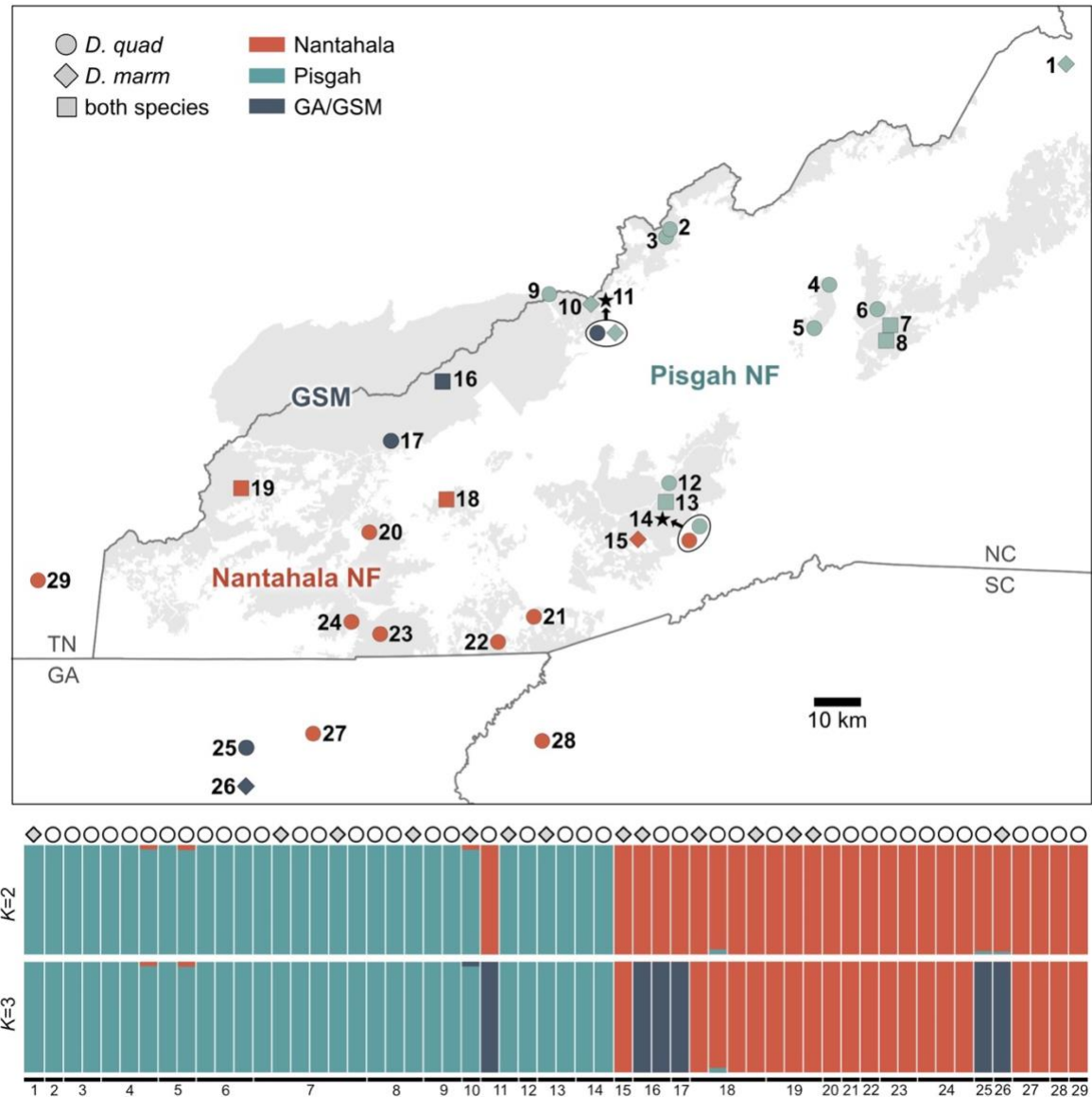


Figure A2. Map of sampling sites and Admixture results for $K = 3$. Top: Map of sampling sites for *D. quadramaculatus*, *D. marmoratus*, and both species within western North Carolina and adjacent states. Numbers on sampling sites correspond to Table A1. Sampling sites are colored by Admixture results for $K = 3$. Forested areas colored light grey. Bottom: Results from Admixture for $K = 2$ and 3 arranged by sampling site, showing individuals grouping by geography rather than phenotype and the subset of individuals in Nantahala grouping separately into a third cluster: Georgia (GA) and Great Smoky Mountains (GSM).

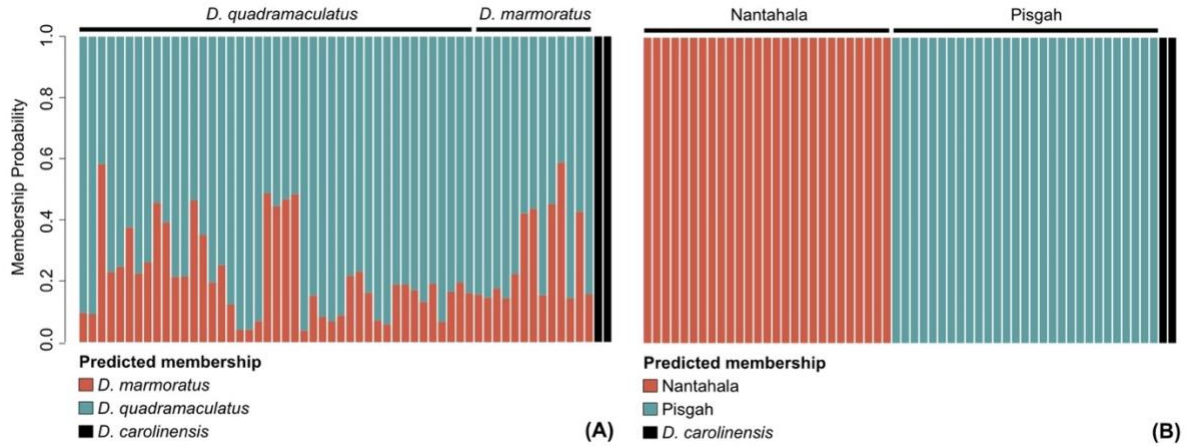


Figure A3. Plots of assignment probabilities from Discriminant Analysis of Principle Components (DAPC). Groups to which individuals were pre-assigned are listed above the plots and colors indicated the predicted membership as determined by DAPC are listed below. (A) Pre-assigned membership by phenotype (*D. quadramaculatus*, *D. marmoratus*, or *D. carolinensis*) resulted in a membership probability of 1.0 for only *D. carolinensis*. DAPC was unable to differentiate between *D. quadramaculatus* and *D. marmoratus* in discriminant space, leading to mixed assignments for all individuals. (B) Pre-assigned membership by geography (Nantahala, Pisgah, or *D. carolinensis* outgroup) correctly predicted membership assignment for all individuals with a membership probability of 1.0.

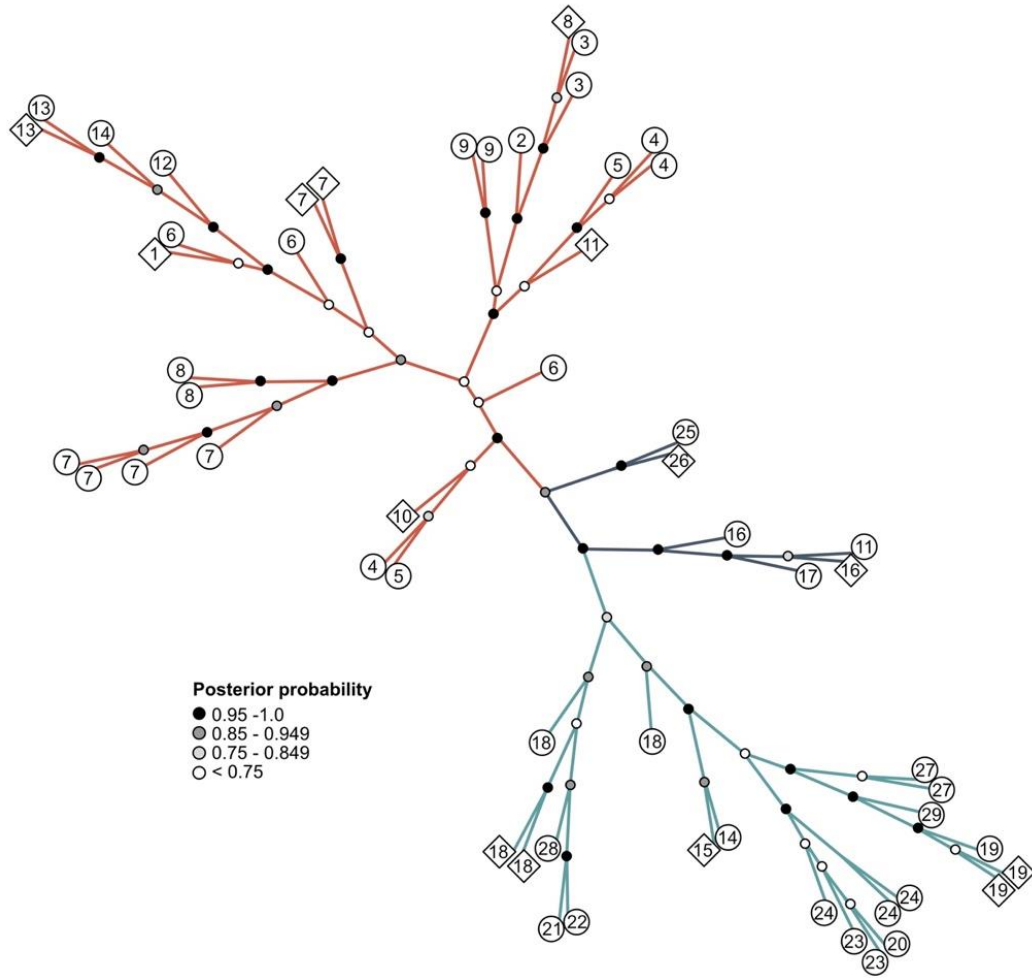


Figure A4. Unrooted SVDquartets tree. *D. quadramaculatus* (circles) and *D. marmoratus* (diamonds). Posterior probabilities are indicated by node color, with darker colors indicating higher support. Colors of branches correspond to Admixture plots for $K = 3$: Nantahala (orange, top); GSM/GA (dark blue, center); and Pisgah (turquoise, bottom).

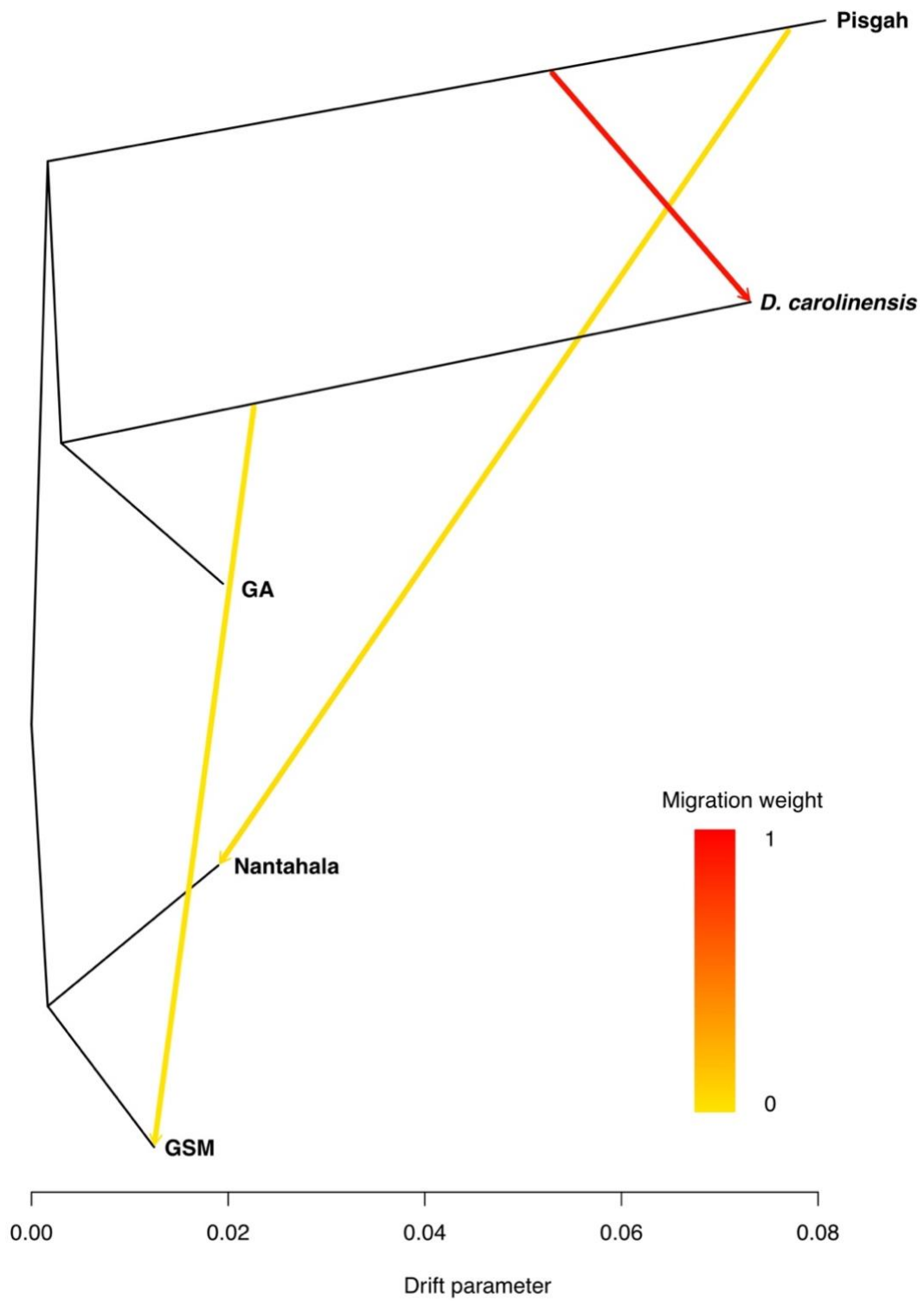


Figure A5. TreeMix plot showing relationships between groups based on the four deepest splits recovered in the BEAST ultrametric tree. Plot shows an unrooted maximum likelihood tree with branch lengths measured by the amount of genetic drift that has occurred between lineages since splitting. Migration edges are shown as colored arrows and indicate possible admixture events between two branches. Higher migration weights indicate greater admixture between the two branches. The results show two low weight migration edges, *D. carolinensis* → GSM and Pisgah → Nantahala, and one strongly weighted migration edge, Pisgah → *D. carolinensis*. Note that arrows do not necessarily indicate unidirectional gene flow and that missing tips may lead to skewed relationships between groups. Hence the results do not necessarily indicate that there was an admixture event directly between Pisgah and *D. carolinensis*, but that an admixture event may have occurred between, for example, Pisgah and a close relative of *D. carolinensis* not included in this study. The results also do not rule out the possibility that relationships have been obscured by incomplete lineage sorting.

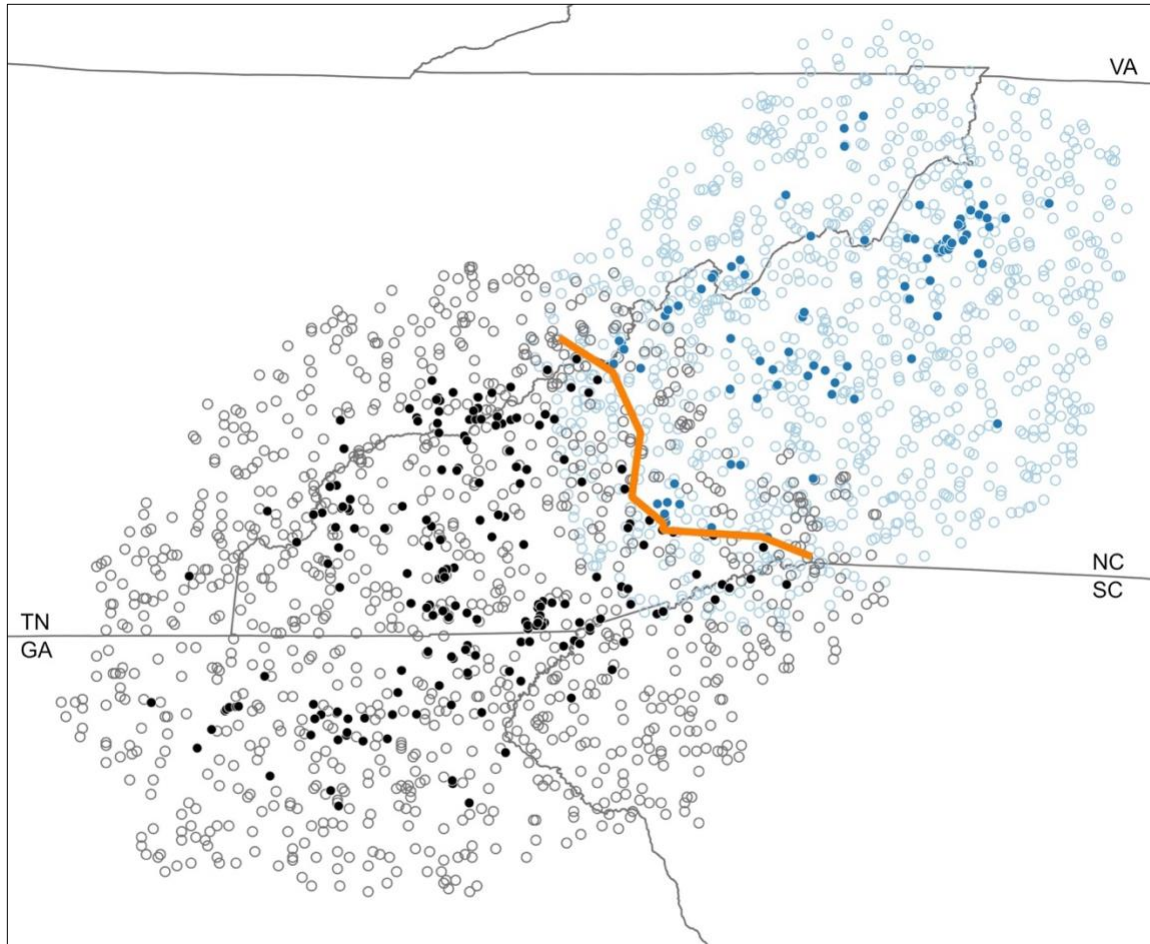


Figure A6. Map of sampling points used for ecological niche models. Presence (occurrence) points and background points randomly selected from a 30 km buffer around presence points. Closed blue circles = Pisgah presence points; open blue circles = Pisgah background points; closed black circles = Nantahala presence points; open black circles = Nantahala background points. The approximate genetic boundary between the two lineages is marked with an orange line. State boundaries are marked with grey lines and labeled with state abbreviations (VA = Virginia; NC = North Carolina; SC = South Carolina; TN = Tennessee; GA = Georgia).

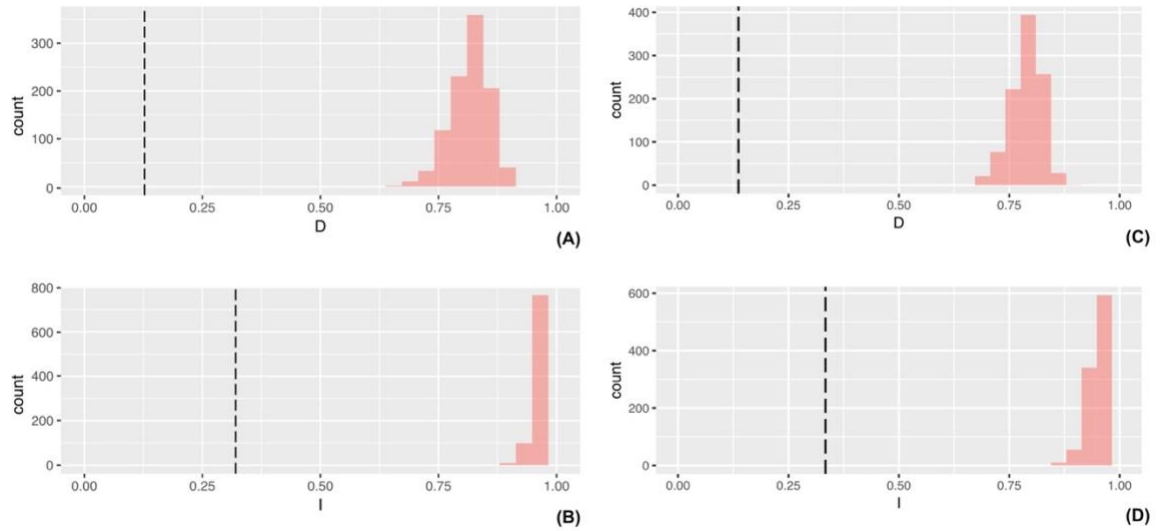


Figure A7. Results of symmetrical background test (A & B) and identity test (C & D) evaluating whether the ecological niche models (ENMs) for Nantahala and Pisgah were significantly different from each other. The dashed line indicates the empirical measure for each test statistic and the pink curve shows the null distribution. Results are considered significant if the empirical measure falls outside of the null distribution, which represents the probability distribution of ENM overlap between Nantahala and Pisgah if occurrences were random. Empirical values for the (A) D and (B) I test statistics ($P < 0.01$) were significantly different than the null distribution for the symmetrical background test, indicating that similarity between the ENMs of both lineages is significantly less than expected. The identity test pools and randomly reassigns presence points to each lineage, building replicate ENMs to produce the null distribution. Empirical values for the (C) D and (D) I test statistics ($P = 0.01$) were significantly different than the null distribution for the identity test, indicating that the distribution of individuals in each lineage is not random with respect to the available climate characteristics.

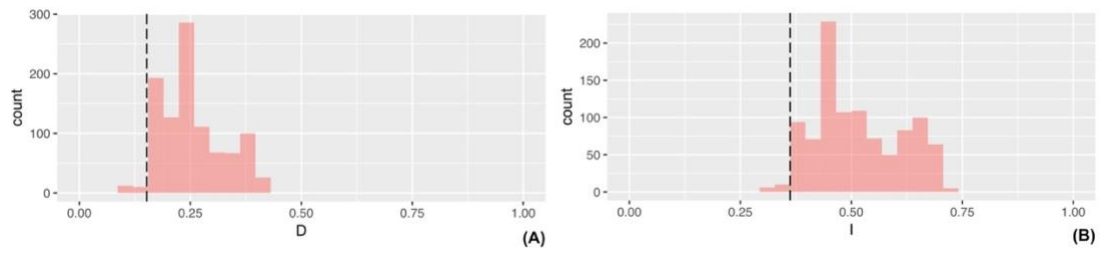


Figure A8. Blob range-break test results for Nantahala versus Pisgah. The dashed line represents the empirical measure for each statistic and the pink curve shows the null distribution (i.e., the probability distribution of ENM overlap between Nantahala and Pisgah if occurrences were random). Empirical values for the **(A)** *D* and **(B)** *I* test statistics ($P < 0.05$) were significantly different than the null distribution, indicating that the climate characteristics significantly differ between these two geographic areas.

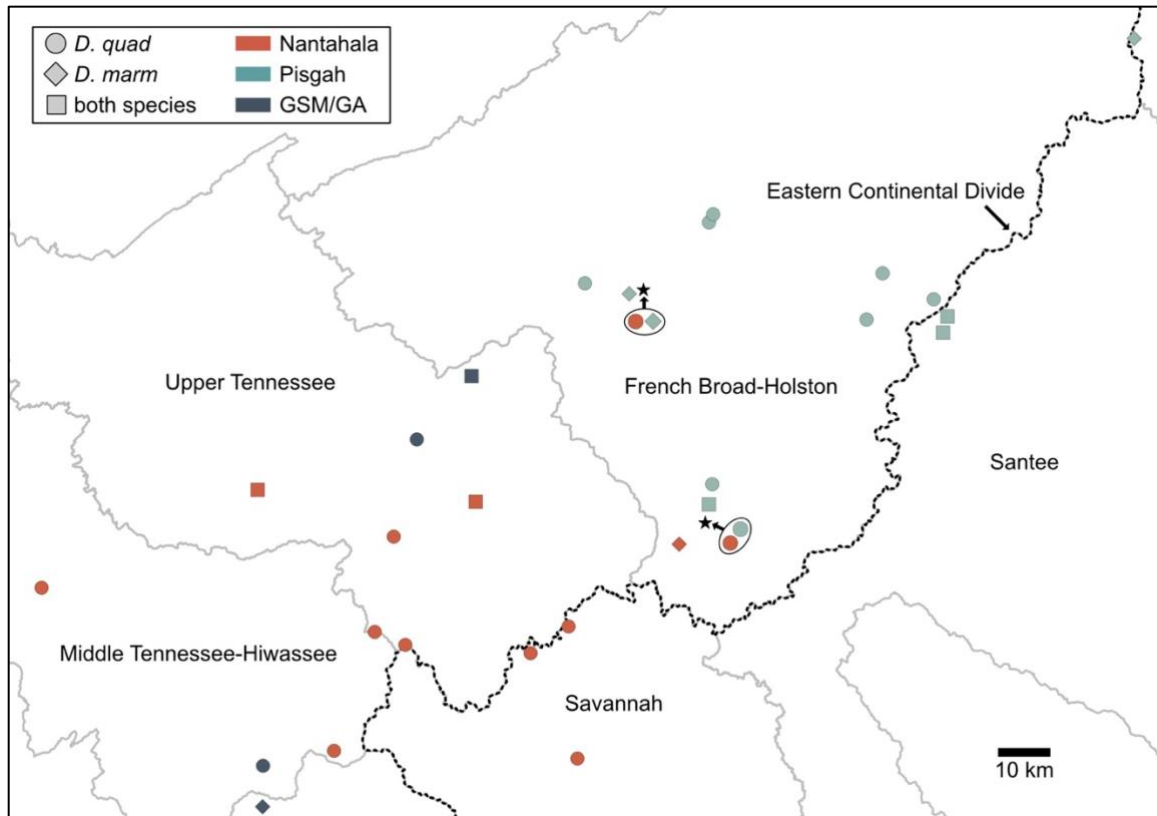


Figure A9. Map showing relationship between major genetic divisions within *D. quadramaculatus* and river basins. The Eastern Continental Divide, shown with a black dashed line, forms the major division between drainage basins in the area. River basins are labeled and outlined in grey. No modern drainage division aligns with the genetic divide found in this study, but it is possible that drainages could have shifted and played a role in past genetic divergence.

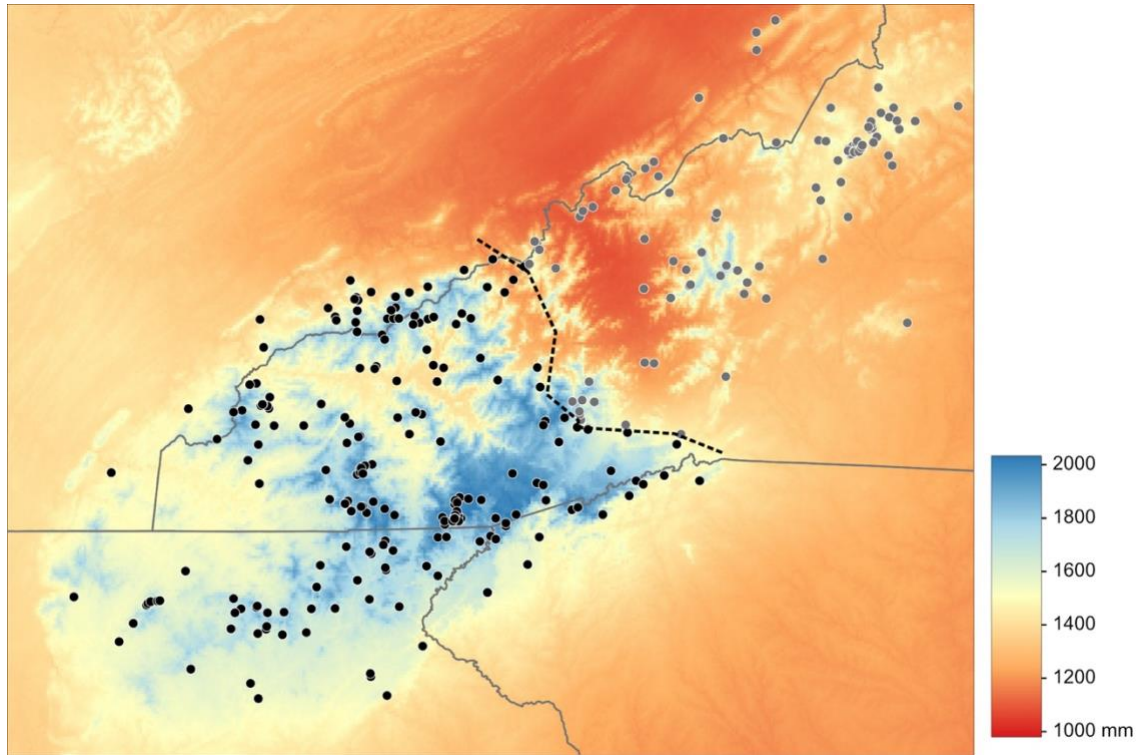


Figure A10. Map of annual precipitation (BIO12) across southern Appalachia. Closed grey circles indicate presence points for Pisgah and closed black circles indicate presence points for Nantahala. Approximate genetic boundary between the two lineages is marked with a dashed line. Annual precipitation values range from 2073 mm to 976 mm per annum across the region. For reference, mean precipitation at presence points was 1712 ± 195 mm for Nantahala and 1433 ± 169 mm for Pisgah.

APPENDIX B. INDIVIDUALS USED IN ANALYSES

Table B1. List of individuals used in analyses for Chapter 3. IDs are the original species identification codes for individuals used in Pyron et al. 2022.

| ID | NCBI Accession | Latitude | Longitude | Clade | Lineage |
|----------|----------------|----------|-----------|-----------|-------------------|
| DAB1271 | SRR16784738 | 34.7678 | -83.9465 | Nantahala | folkertsi |
| DAB14959 | SRR16762478 | 34.7678 | -83.9465 | Nantahala | folkertsi |
| DAB1815 | SRR16784516 | 34.6675 | -83.3163 | Nantahala | folkertsi |
| DAB6574 | SRR16784316 | 34.86826 | -83.80813 | Nantahala | folkertsi |
| DAB7282 | SRR16784291 | 34.55828 | -84.25329 | Nantahala | folkertsi |
| DAB9079 | SRR16784723 | 35.02021 | -83.55402 | Nantahala | folkertsi |
| DAB10796 | SRR16784315 | 34.70757 | -83.91559 | Nantahala | marmoratus B |
| DAB10834 | SRR16784304 | 34.95194 | -83.49281 | Nantahala | marmoratus B |
| DAB10839 | SRR16820807 | 34.80818 | -83.59244 | Nantahala | marmoratus B |
| DAB10866 | SRR16784293 | 34.77672 | -83.73888 | Nantahala | marmoratus B |
| DAB10879 | SRR16784282 | 35.02087 | -83.11401 | Nantahala | marmoratus B |
| DAB1336 | SRR16784472 | 34.9804 | -83.1477 | Nantahala | marmoratus B |
| DAB9084 | SRR16784240 | 35.02021 | -83.55402 | Nantahala | marmoratus B |
| DAB9281 | SRR16784232 | 34.83089 | -83.62661 | Nantahala | marmoratus B |
| DAB10816 | SRR16784186 | 34.85984 | -83.38285 | Nantahala | quadramaculatus A |
| DAB11714 | SRR16784161 | 35.32147 | -82.84789 | Nantahala | quadramaculatus A |
| DAB11971 | SRR16784151 | 35.2007 | -82.6438 | Nantahala | quadramaculatus A |
| DAB12008 | SRR16784148 | 35.2572 | -83.22331 | Nantahala | quadramaculatus A |
| DAB12028 | SRR16784565 | 35.19353 | -82.97256 | Nantahala | quadramaculatus A |
| DAB12056 | SRR16784146 | 35.2738 | -83.00031 | Nantahala | quadramaculatus A |
| DAB12058 | SRR16784563 | 35.27982 | -82.9669 | Nantahala | quadramaculatus A |
| DAB12170 | SRR16784136 | 35.32443 | -84.17772 | Nantahala | quadramaculatus A |
| DAB12636 | SRR16784548 | 34.95121 | -82.63043 | Nantahala | quadramaculatus A |
| DAB1338 | SRR16784727 | 34.9804 | -83.1477 | Nantahala | quadramaculatus A |
| DAB4678 | SRR16784401 | 34.67148 | -83.73616 | Nantahala | quadramaculatus A |
| DAB4860 | SRR16784389 | 34.64158 | -83.94225 | Nantahala | quadramaculatus A |
| DAB5148 | SRR16784372 | 35.13442 | -83.61669 | Nantahala | quadramaculatus A |
| DAB5179 | SRR16784369 | 35.1596 | -82.97338 | Nantahala | quadramaculatus A |
| DAB5385 | SRR16784364 | 35.0672 | -83.5975 | Nantahala | quadramaculatus A |
| DAB5482 | SRR16820805 | 35.19924 | -83.93872 | Nantahala | quadramaculatus A |
| DAB5507 | SRR16784355 | 35.11796 | -83.83916 | Nantahala | quadramaculatus A |
| DAB5553 | SRR16820804 | 35.07165 | -83.22971 | Nantahala | quadramaculatus A |
| DAB5947 | SRR16762499 | 34.19131 | -83.12651 | Nantahala | quadramaculatus A |
| DAB6549 | SRR16784318 | 34.9772 | -83.30237 | Nantahala | quadramaculatus A |

Table B1 (continued)

| | | | | | |
|----------|-------------|----------|-----------|-----------|-------------------|
| DAB6558 | SRR16784317 | 35.07953 | -83.14342 | Nantahala | quadramaculatus A |
| DAB6843 | SRR16784306 | 35.27679 | -82.31764 | Nantahala | quadramaculatus A |
| DAB7273 | SRR16784294 | 35.21898 | -84.03545 | Nantahala | quadramaculatus A |
| DAB7495 | SRR16784289 | 34.73068 | -83.38515 | Nantahala | quadramaculatus A |
| DAB8057 | SRR16784275 | 35.06146 | -84.16353 | Nantahala | quadramaculatus A |
| DAB8081 | SRR16784273 | 35.20937 | -83.45065 | Nantahala | quadramaculatus A |
| DAB8128 | SRR16784360 | 35.42278 | -83.19018 | Nantahala | quadramaculatus A |
| DAB9112 | SRR16784237 | 34.61559 | -84.19755 | Nantahala | quadramaculatus A |
| DAB9158 | SRR16784718 | 34.70714 | -83.98533 | Nantahala | quadramaculatus A |
| DAB9219 | SRR16784234 | 35.42871 | -83.88972 | Nantahala | quadramaculatus A |
| DAB9395 | SRR16784703 | 35.11649 | -83.54473 | Nantahala | quadramaculatus A |
| DAB9421 | SRR16784224 | 35.07023 | -83.49728 | Nantahala | quadramaculatus A |
| RAP0678 | SRR16762488 | 34.777 | -84.33188 | Nantahala | quadramaculatus A |
| DAB14845 | SRR16762480 | 35.54213 | -83.11921 | Nantahala | quadramaculatus F |
| DAB4058 | SRR16784427 | 35.63478 | -83.49663 | Nantahala | quadramaculatus F |
| DAB4134 | SRR16784423 | 35.68247 | -83.63821 | Nantahala | quadramaculatus F |
| DAB8944 | SRR16784246 | 35.49008 | -83.1642 | Nantahala | quadramaculatus F |
| RAP0839 | SRR16762487 | 35.55528 | -83.06935 | Nantahala | quadramaculatus F |
| RAP0888 | SRR16762476 | 35.63002 | -83.21524 | Nantahala | quadramaculatus F |
| DAB5648 | SRR16784449 | 35.34768 | -83.97356 | Pisgah | marmoratus C |
| DAB10780 | SRR16784187 | 35.26575 | -83.68887 | Pisgah | marmoratus C |
| DAB10870 | SRR16820806 | 35.06901 | -83.62981 | Pisgah | marmoratus C |
| RAP1064 | SRR16684300 | 35.26449 | -83.583 | Pisgah | marmoratus C |
| DAB4928 | SRR16784404 | 35.3347 | -83.3722 | Pisgah | marmoratus C |
| DAB8129 | SRR16784116 | 35.42278 | -83.19018 | Pisgah | marmoratus C |
| DAB8931 | SRR16784326 | 35.49008 | -83.1642 | Pisgah | marmoratus C |
| DAB14829 | SRR16762481 | 35.43366 | -83.09119 | Pisgah | marmoratus C |
| DAB6085 | SRR16784561 | 35.73787 | -83.01598 | Pisgah | marmoratus C |
| RAP0999 | SRR16684312 | 35.46852 | -82.95499 | Pisgah | marmoratus C |
| DAB12175 | SRR16784135 | 35.85946 | -82.90981 | Pisgah | marmoratus C |
| RAP0985 | SRR16684314 | 35.32088 | -82.85289 | Pisgah | marmoratus C |
| DAB14609 | SRR16762509 | 35.42821 | -82.83357 | Pisgah | marmoratus C |
| RAP0995 | SRR16684313 | 35.42858 | -82.83357 | Pisgah | marmoratus C |
| RAP1057 | SRR16684304 | 35.99453 | -82.70462 | Pisgah | marmoratus C |
| DAB13566 | SRR16784510 | 36.02739 | -82.61453 | Pisgah | marmoratus C |
| DAB7837 | SRR16784753 | 35.74139 | -83.05033 | Pisgah | quadramaculatus C |
| DAB10487 | SRR16784198 | 35.78168 | -83.0274 | Pisgah | quadramaculatus C |

Table B1 (continued)

| | | | | | |
|----------|-------------|----------|-----------|--------|-------------------|
| DAB6079 | SRR16784382 | 35.73787 | -83.01598 | Pisgah | quadramaculatus C |
| RAP0824 | SRR16684319 | 35.74365 | -83.00972 | Pisgah | quadramaculatus C |
| DAB10504 | SRR16784197 | 35.79891 | -82.95174 | Pisgah | quadramaculatus C |
| DAB12176 | SRR16784134 | 35.85946 | -82.90981 | Pisgah | quadramaculatus C |
| DAB1355 | SRR16784705 | 35.6875 | -82.8971 | Pisgah | quadramaculatus C |
| DAB9441 | SRR16784222 | 35.94122 | -82.892 | Pisgah | quadramaculatus C |
| DAB9400 | SRR16784702 | 35.96998 | -82.85129 | Pisgah | quadramaculatus C |
| RAP1060 | SRR16684301 | 35.83053 | -82.84617 | Pisgah | quadramaculatus C |
| DAB10678 | SRR16784190 | 35.92808 | -82.80268 | Pisgah | quadramaculatus C |
| DAB10666 | SRR16784624 | 35.95079 | -82.78965 | Pisgah | quadramaculatus C |
| DAB10555 | SRR16784195 | 36.01567 | -82.73541 | Pisgah | quadramaculatus C |
| DAB10577 | SRR16784629 | 36.01567 | -82.73541 | Pisgah | quadramaculatus C |
| RAP1058 | SRR16684303 | 35.99453 | -82.70462 | Pisgah | quadramaculatus C |
| DAB13565 | SRR16784511 | 36.02739 | -82.61453 | Pisgah | quadramaculatus C |
| DAB10688 | SRR16784189 | 36.03441 | -82.59556 | Pisgah | quadramaculatus C |
| RAP0804 | SRR16684320 | 35.7233 | -82.40522 | Pisgah | quadramaculatus C |
| DAB4956 | SRR16784378 | 35.7091 | -82.39394 | Pisgah | quadramaculatus C |
| DAB13556 | SRR16784513 | 35.8036 | -82.35425 | Pisgah | quadramaculatus C |
| DAB13559 | SRR16784512 | 35.8036 | -82.35425 | Pisgah | quadramaculatus C |
| RAP1052 | SRR16684307 | 36.25141 | -82.36431 | Pisgah | quadramaculatus G |
| RAP1103 | SRR16684288 | 36.19457 | -82.34739 | Pisgah | quadramaculatus G |
| DAB11014 | SRR16784176 | 36.19263 | -82.3387 | Pisgah | quadramaculatus G |
| DAB10922 | SRR16784183 | 36.12579 | -82.31167 | Pisgah | quadramaculatus G |
| DAB10950 | SRR16784180 | 36.1292 | -82.26374 | Pisgah | quadramaculatus G |
| DAB11798 | SRR16784158 | 36.14873 | -82.16458 | Pisgah | quadramaculatus G |
| DAB11836 | SRR16784155 | 36.21979 | -82.08759 | Pisgah | quadramaculatus G |
| DAB7914 | SRR16784280 | 36.12162 | -82.08555 | Pisgah | quadramaculatus G |
| DAB7888 | SRR16784283 | 36.18089 | -82.01077 | Pisgah | quadramaculatus G |
| DAB8408 | SRR16784337 | 36.04009 | -82.37661 | Pisgah | marmoratus E |
| RAP1097 | SRR16684289 | 36.25186 | -82.36518 | Pisgah | marmoratus E |
| RAP0762 | SRR16684275 | 36.08215 | -82.27927 | Pisgah | marmoratus E |
| DAB8739 | SRR16784254 | 35.75402 | -82.10065 | Pisgah | marmoratus E |
| DAB8289 | SRR16820809 | 35.74343 | -82.05918 | Pisgah | marmoratus E |
| RAP0740 | SRR16684279 | 36.10743 | -81.78184 | Pisgah | marmoratus E |
| DAB14651 | SRR16762514 | 35.60436 | -81.78175 | Pisgah | marmoratus E |
| RAP1006 | SRR16684311 | 35.60424 | -81.7817 | Pisgah | marmoratus E |
| DAB13291 | SRR16784515 | 36.65481 | -81.58453 | Pisgah | marmoratus E |

Table B1 (continued)

| | | | | | |
|----------|-------------|----------|-----------|--------|-------------------|
| DAB6723 | SRR16784778 | 35.37239 | -82.78314 | Pisgah | quadramaculatus E |
| DAB11796 | SRR16784159 | 35.2856 | -82.7426 | Pisgah | quadramaculatus E |
| DAB955 | SRR16784127 | 35.26765 | -82.72826 | Pisgah | quadramaculatus E |
| DAB6073 | SRR16784343 | 35.50086 | -82.61127 | Pisgah | quadramaculatus E |
| DAB10560 | SRR16784194 | 35.96414 | -82.5882 | Pisgah | quadramaculatus E |
| DAB10591 | SRR16784628 | 35.96414 | -82.5882 | Pisgah | quadramaculatus E |
| RAP1055 | SRR16684306 | 36.08279 | -82.50627 | Pisgah | quadramaculatus E |
| DAB8335 | SRR16784267 | 36.05402 | -82.42011 | Pisgah | quadramaculatus E |
| DAB8405 | SRR16784349 | 36.04009 | -82.37661 | Pisgah | quadramaculatus E |
| DAB1080 | SRR16784760 | 35.47187 | -82.36341 | Pisgah | quadramaculatus E |
| RAP0757 | SRR16684276 | 36.11097 | -82.35898 | Pisgah | quadramaculatus E |
| DAB13647 | SRR16784506 | 35.92331 | -82.30004 | Pisgah | quadramaculatus E |
| RAP0763 | SRR16684324 | 36.08215 | -82.27927 | Pisgah | quadramaculatus E |
| RAP0800 | SRR16684321 | 35.63106 | -82.26091 | Pisgah | quadramaculatus E |
| RAP0772 | SRR16684323 | 36.13628 | -82.23947 | Pisgah | quadramaculatus E |
| DAB6345 | SRR16784323 | 35.74073 | -82.19735 | Pisgah | quadramaculatus E |
| RAP0640 | SRR16684277 | 35.75641 | -82.14823 | Pisgah | quadramaculatus E |
| DAB8741 | SRR16784253 | 35.75402 | -82.10065 | Pisgah | quadramaculatus E |
| DAB11826 | SRR16784580 | 36.11128 | -82.06542 | Pisgah | quadramaculatus E |
| DAB11820 | SRR16784156 | 36.37031 | -82.05783 | Pisgah | quadramaculatus E |
| DAB11869 | SRR16784154 | 36.50857 | -81.98661 | Pisgah | quadramaculatus E |
| DAB8727 | SRR16784255 | 35.81198 | -81.98038 | Pisgah | quadramaculatus E |
| DAB7139 | SRR16784297 | 35.95109 | -81.94415 | Pisgah | quadramaculatus E |
| RAP0787 | SRR16684322 | 35.81743 | -81.93889 | Pisgah | quadramaculatus E |
| DAB10992 | SRR16784177 | 36.1941 | -81.86932 | Pisgah | quadramaculatus E |
| DAB11022 | SRR16784174 | 36.13241 | -81.8692 | Pisgah | quadramaculatus E |
| RAP0638 | SRR16684278 | 36.04457 | -81.84618 | Pisgah | quadramaculatus E |
| DAB7078 | SRR16784300 | 36.01184 | -81.7374 | Pisgah | quadramaculatus E |
| DAB8413 | SRR16784264 | 36.27119 | -81.71533 | Pisgah | quadramaculatus E |
| DAB8238 | SRR16784268 | 35.64383 | -81.70957 | Pisgah | quadramaculatus E |
| RAP0730 | SRR16684280 | 36.22365 | -81.66618 | Pisgah | quadramaculatus E |
| RAP0716 | SRR16684282 | 36.55551 | -81.64675 | Pisgah | quadramaculatus E |
| DAB11004 | SRR16784609 | 36.15129 | -81.63065 | Pisgah | quadramaculatus E |
| DAB12300 | SRR16784126 | 36.42332 | -81.59531 | Pisgah | quadramaculatus E |
| DAB7096 | SRR16784299 | 36.04851 | -81.59473 | Pisgah | quadramaculatus E |
| RAP0720 | SRR16684281 | 36.42378 | -81.59439 | Pisgah | quadramaculatus E |
| DAB10411 | SRR16784202 | 36.39178 | -81.40398 | Pisgah | quadramaculatus E |

Table B1 (continued)

| | | | | | |
|----------|-------------|----------|-----------|--------|-------------------|
| DAB10279 | SRR16784654 | 36.37058 | -81.39512 | Pisgah | quadramaculatus E |
| DAB11904 | SRR16784152 | 36.76247 | -81.49227 | Pisgah | quadramaculatus D |
| DAB10383 | SRR16784606 | 36.31753 | -81.25901 | Pisgah | quadramaculatus D |
| DAB10230 | SRR16784207 | 36.33681 | -81.16466 | Pisgah | quadramaculatus D |
| DAB10100 | SRR16784679 | 36.30539 | -81.1634 | Pisgah | quadramaculatus D |
| RAP0948 | SRR16684315 | 36.37189 | -81.15719 | Pisgah | quadramaculatus D |
| DAB1469 | SRR16784638 | 36.0952 | -81.08887 | Pisgah | quadramaculatus D |
| DAB10175 | SRR16784614 | 36.42095 | -80.93005 | Pisgah | quadramaculatus D |
| DAB10214 | SRR16784208 | 36.44749 | -80.84788 | Pisgah | quadramaculatus D |
| DAB3262 | SRR16784452 | 36.6121 | -80.77138 | Pisgah | quadramaculatus D |
| DAB3390 | SRR16784447 | 37.43269 | -80.51239 | Pisgah | quadramaculatus D |
| DAB11302 | SRR16784163 | 36.50787 | -80.48754 | Pisgah | quadramaculatus D |
| DAB11105 | SRR16784172 | 36.6987 | -80.44647 | Pisgah | quadramaculatus D |
| DAB9597 | SRR16784691 | 36.38027 | -80.18525 | Pisgah | quadramaculatus D |

REFERENCES

- Aguillon, S. M., J. W. Fitzpatrick, R. Bowman, S. J. Schoech, A. G. Clark, G. Coop, and N. Chen. 2017. Deconstructing isolation-by-distance: The genomic consequences of limited dispersal. *PLoS Genet* 13:1–27.
- Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela, and R. P. Anderson. 2015. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38:541–545.
- Alexander, D. H., and K. Lange. 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12:246.
- Allen, K. E., W. P. Tapondjou, B. Freeman, J. C. Cooper, R. M. Brown, and A. T. Peterson. 2021. Modelling potential Pleistocene habitat corridors between Afromontane forest regions. *Biodivers Conserv* 30:2361–2375.
- Altman, D. G., and J. M. Bland. 1995. Statistics notes: Absence of evidence is not evidence of absence. *BMJ* 311:485–485.
- Alvarado-Serrano, D. F., and L. L. Knowles. 2014. Ecological niche models in phylogeographic studies: Applications, advances and precautions. *Mol Ecol Resour* 14:233–248.
- Anderson, R. P., and I. Gonzalez. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecol Modell* 222:2796–2811.
- Andrews, K. R., and G. Luikart. 2014. Recent novel approaches for population genomics data analysis. *Mol Ecol* 23:1661–1667.
- Antunes, B., C. Figueiredo-Vázquez, K. Dudek, M. Liana, M. Pabijan, P. Zieliński, and W. Babik. 2022. Landscape genetics reveals contrasting patterns of connectivity in two newt species (*Lissotriton montandoni* and *L. vulgaris*). *Mol Ecol* 1–16.
- Araújo, M. B., D. Nogués-Bravo, J. A. F. Diniz-Filho, A. M. Haywood, P. J. Valdes, and C. Rahbek. 2008. Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography* 31:8–15.
- Årevall, J., R. Early, A. Estrada, U. Wennergren, and A. C. Eklöf. 2018. Conditions for successful range shifts under climate change: The role of species dispersal and landscape configuration. *Divers Distrib* 24:1598–1611.
- Arnaud, J. F. 2003. Metapopulation genetic structure and migration pathways in the land snail *Helix aspersa*: Influence of landscape heterogeneity. *Landsc Ecol* 18:333–346.
- Arnégard, M. E., M. D. McGee, B. Matthews, K. B. Marchinko, G. L. Conte, S. Kabir, N. Bedford, S. Bergek, Y. F. Chan, F. C. Jones, D. M. Kingsley, C. L. Peichel, and D. Schluter. 2014. Genetics of ecological divergence during speciation. *Nature* 511:307–311.
- Arnold, M. L., and K. Kunte. 2017. Adaptive Genetic Exchange: A Tangled History of Admixture and Evolutionary Innovation. *Trends Ecol Evol* 32:601–611. Elsevier Ltd.
- Ash, A. N. 1997. Disappearance and return of plethodontid salamanders to clearcut plots in the southern Blue Ridge Mountains. *Conservation Biology* 11:983–989.

- Ashraf, B., and D. J. Lawson. 2021. Genetic drift from the out-of-Africa bottleneck leads to biased estimation of genetic architecture and selection. *European Journal of Human Genetics* 29:1549–1556.
- Ashraf, U., A. T. Peterson, M. N. Chaudhry, I. Ashraf, Z. Saqib, S. R. Ahmad, and H. Ali. 2017. Ecological niche model comparison under different climate scenarios: A case study of *Olea* spp. in Asia. *Ecosphere* 8:1–13.
- Avice, J. C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annu Rev Ecol Syst* 18:489–522.
- Avice, J. C. 2009. Phylogeography: Retrospect and prospect. *J Biogeogr* 36:3–15.
- Avice, J. C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press.
- Bailey, S. F., N. Rodrigue, and R. Kassen. 2015. The effect of selection environment on the probability of parallel evolution. *Mol Biol Evol* 32:1436–1448.
- Bandelt, H. J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48.
- Barbet-Massin, M., and W. Jetz. 2014. A 40-year, continent-wide, multispecies assessment of relevant climate predictors for species distribution modelling. *Divers Distrib* 20:1285–1295.
- Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol Evol* 3:327–338.
- Barnard-Kubow, K. B., C. L. Debban, and L. F. Galloway. 2015. Multiple glacial refugia lead to genetic structuring and the potential for reproductive isolation in a herbaceous plant. *Am J Bot* 102:1842–1853.
- Barnett, J. B., C. Michalis, N. E. Scott-Samuel, and I. C. Cuthill. 2021. Colour pattern variation forms local background matching camouflage in a leaf-mimicking toad. *J Evol Biol* 34:1531–1540.
- Barrett, K., B. S. Helms, C. Guyer, and J. E. Schoonover. 2010. Linking process to pattern: Causes of stream-breeding amphibian decline in urbanized watersheds. *Biol Conserv* 143:1998–2005.
- Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberón, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol Modell* 222:1810–1819.
- Beachy, C., and R. Bruce. 2003. Life history of a small form of the plethodontid salamander *Desmognathus quadramaculatus*. *Amphibia-Reptilia* 24:13–26.
- Beamer, D. A., and T. Lamb. 2008. Dusky salamanders (*Desmognathus*, Plethodontidae) from the Coastal Plain: Multiple independent lineages and their bearing on the molecular phylogeny of the genus. *Mol Phylogenet Evol* 47:143–153.
- Beamer, D. A., and T. Lamb. 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.
- Beheregaray, L. B. 2008. Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. *Mol Ecol* 17:3754–3774.

- Bemmels, J. B., and C. W. Dick. 2018. Genomic evidence of a widespread southern distribution during the Last Glacial Maximum for two eastern North American hickory species. *J Biogeogr* 45:1739–1750.
- Bernardo, J., and J. R. Spotila. 2006. Physiological constraints on organismal response to global warming: mechanistic insights from clinally varying populations and implications for assessing endangerment. *Biol Lett* 2:135–139.
- Berner, D., and W. Salzburger. 2015. The genomics of organismal diversification illuminated by adaptive radiations. *Trends in Genetics* 31:491–499.
- Bhatia, G., N. Patterson, S. Sankararaman, and A. L. Price. 2013. Estimating and interpreting FST: The impact of rare variants. *Genome Res* 23:1514–1521.
- Bickford, D., D. J. Lohman, N. S. Sodhi, P. K. L. Ng, R. Meier, K. Winker, K. K. Ingram, and I. Das. 2007. Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155.
- Bliss, M., and K. K. Cecala. 2016. Does habitat disturbance affect the behaviors of Appalachian stream salamanders? *Herpetol Conserv Biol* 10:811–818.
- Bobrowski, M., J. Weidinger, and U. Schickhoff. 2021. Is new always better? Frontiers in global climate datasets for modeling treeline species in the Himalayas. *Atmosphere (Basel)* 12:543.
- Bohl, C. L., J. M. Kass, and R. P. Anderson. 2019. A new null model approach to quantify performance and significance for ecological niche models of species distributions. *J Biogeogr* 46:1101–1111.
- Bolnick, D. I., and B. M. Fitzpatrick. 2007. Sympatric speciation: Models and empirical evidence. *Annu Rev Ecol Evol Syst* 38:459–487.
- Davis, M. B. 1983. Quaternary History of Deciduous Forests of Eastern North America and Europe. *Annals of the Missouri Botanical Garden* 70:550–563.
- Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 10:1–6.
- Bouزيد, N. M., J. W. Archie, R. A. Anderson, J. A. Grummer, and A. D. Leaché. 2022. Evidence for ephemeral ring species formation during the diversification history of western fence lizards (*Sceloporus occidentalis*). *Mol Ecol* 31:620–631.
- Bowler, D. E., C. T. Callaghan, N. Bhandari, K. Henle, M. Benjamin Barth, C. Koppitz, R. Klenke, M. Winter, F. Jansen, H. Bruelheide, and A. Bonn. 2022. Temporal trends in the spatial bias of species occurrence records. *Ecography* 2022:1–13.
- Branch, C. L., J. P. Jahner, D. Y. Kozlovsky, T. L. Parchman, and V. V Pravosudov. 2017. Absence of population structure across elevational gradients despite large phenotypic variation in mountain chickadees (*Poecile gambeli*). *R Soc Open Sci* 4:170057.
- Brock, C. D., M. E. Cummings, and D. I. Bolnick. 2017. Phenotypic plasticity drives a depth gradient in male conspicuousness in threespine stickleback, *Gasterosteus aculeatus*. *Evolution (N Y)* 71:2022–2036.
- Broennimann, O., M. C. Fitzpatrick, P. B. Pearman, B. Petitpierre, L. Pellissier, N. G. Yoccoz, W. Thuiller, M. J. Fortin, C. Randin, N. E. Zimmermann, C. H. Graham, and A. Guisan. 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography* 21:481–497.

- Bruce, R. C. 2011. Community Assembly in the Salamander Genus *Desmognathus*. *Herpetological Monographs* 25:1–24.
- Bruce, R. C. 1985. Larval periods, population structure and the effects of stream drift in larvae of the salamanders *Desmognathus quadramaculatus* and *Leurognathus marmoratus* in a southern Appalachian stream. *Copeia* 1985:847–854.
- Bruce, R. C. 1988. Life history variation in the salamander *Desmognathus quadramaculatus*. *Herpetologica* 44:218–227.
- Burgon, J. D., D. R. Vieites, A. Jacobs, S. K. Weidt, H. M. Gunter, S. Steinfartz, K. Burgess, B. K. Mable, and K. R. Elmer. 2020. Functional colour genes and signals of selection in colour-polymorphic salamanders. *Mol Ecol* 29:1284–1299.
- Camp, C. D., J. L. Marshall, and R. M. Austin, Jr. 2000. The evolution of adult body size in black-bellied salamanders (*Desmognathus quadramaculatus* complex). *Can J Zool* 78:1712–1722.
- Camp, C. D., Z. L. Seymour, and J. A. Wooten. 2013a. Morphological variation in the cryptic species *Desmognathus quadramaculatus* (Black-bellied Salamander) and *Desmognathus folkertsi* (Dwarf Black-bellied Salamander). *J Herpetol* 47:471–479.
- Camp, C. D., S. G. Tilley, R. M. Austin, and J. L. Marshall. 2002. A new species of black-bellied salamander (genus *Desmognathus*) from the Appalachian Mountains of northern Georgia. *Herpetologica* 58:471–484.
- Camp, C. D., and J. A. Wooten. 2016. Hidden in plain sight: Cryptic diversity in the Plethodontidae. *Copeia* 2016:111–117.
- Camp, C. D., J. A. Wooten, C. M. Corbet, E. A. Dulka, J. A. Mitchem, and T. J. Krieger. 2013b. Ecological interactions between two broadly sympatric, cryptic species of dusky salamander (Genus *Desmognathus*). *Copeia* 2013:499–506.
- Carlo, M., J. C. Graph, D. Posada, K. A. Crandall, and D. Posada. 2001. Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol Evol* 16:37–45.
- Cayuela, H., Q. Rougemont, J. G. Prunier, J. S. Moore, J. Clobert, A. Besnard, and L. Bernatchez. 2018. Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Mol Ecol* 27:3976–4010.
- Cecala, K. K., W. H. Lowe, and J. C. Maerz. 2014. Riparian disturbance restricts in-stream movement of salamanders. *Freshw Biol* 59:2354–2364.
- Cecala, K. K., J. C. Maerz, B. J. Halstead, J. R. Frisch, T. L. Gragson, J. Hepinstall-Cymerman, D. S. Leigh, C. R. Jackson, J. T. Peterson, and C. M. Pringle. 2018. Multiple drivers, scales, and interactions influence southern Appalachian stream salamander occupancy. *Ecosphere* 9:e02150.
- Ceccarelli, F. S., A. A. Ojanguren-Affilastro, M. J. Ramírez, J. A. Ochoa, C. I. Mattoni, and L. Prendini. 2016. Andean uplift drives diversification of the bothriurid scorpion genus *Brachistosternus*. *J Biogeogr* 43:1942–1954.
- Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890.
- Chernomor, O., A. von Haeseler, and B. Q. Minh. 2016. Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Syst Biol* 65:997–1008.
- Chifman, J., and L. Kubatko. 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30:3317–3324.

- Condamine, F. L., A. Antonelli, L. P. Lagomarsino, C. Hoorn, and L. H. Liow. 2018. Teasing Apart Mountain Uplift, Climate Change and Biotic Drivers of Species Diversification. P. 257 in *Mountains, Climate and Biodiversity*. John Wiley & Sons.
- Coster, S. S., A. B. Welsh, G. Costanzo, S. R. Harding, J. T. Anderson, S. B. McRae, and T. E. Katzner. 2018. Genetic analyses reveal cryptic introgression in secretive marsh bird populations. *Ecol Evol* 8:9870–9879.
- Crespi, E. J., L. J. Rissler, and R. A. Browne. 2003. Testing Pleistocene refugia theory: Phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. *Mol Ecol* 12:969–984.
- Currinder, B., K. K. Cecala, R. M. Northington, and M. E. Dorcas. 2014. Response of stream salamanders to experimental drought in the southern Appalachian Mountains, USA. *J Freshw Ecol* 29:579–587.
- Dagnino, D., L. Minuto, and G. Casazza. 2017. Divergence is not enough: the use of ecological niche models for the validation of taxon boundaries. *Plant Biol* 19:1003–1011.
- Danecek, P., J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, and H. Li. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:1–4.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: More models, new heuristics and parallel computing. *Nat Methods* 9:772–772.
- Darwell, C. T., and J. M. Cook. 2017. Cryptic diversity in a fig wasp community-morphologically differentiated species are sympatric but cryptic species are parapatric. *Mol Ecol* 26:937–950.
- Davis, M. B., and R. G. Shaw. 2001. Range shifts and adaptive responses to quaternary climate change. *Science* (1979) 292:673–679.
- De Jong, G. 2005. Evolution of phenotypic plasticity: Patterns of plasticity and the emergence of ecotypes. *New Phytologist* 166:101–118.
- De La Torre, A. R., D. R. Roberts, and S. N. Aitken. 2014. Genome-wide admixture and ecological niche modelling reveal the maintenance of species boundaries despite long history of interspecific gene flow. *Mol Ecol* 23:2046–2059.
- Degnan, J. H., and N. A. Rosenberg. 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet* 2:762–768.
- Delcourt, H. R., and P. A. Delcourt. 1988. Quaternary landscape ecology: Relevant scales in space and time. *Landsc Ecol* 2:23–44.
- Delcourt, P. A., P. Haccou, and H. R. Delcourt. 2004. *Prehistoric Native Americans and ecological change: human ecosystems in eastern North America since the Pleistocene*. Cambridge University Press.
- della Croce, P., G. C. Poole, and G. Luikart. 2016. Detecting and quantifying introgression in hybridized populations: simplifying assumptions yield overconfidence and uncertainty. *Mol Ecol Resour* 16:1287–1302.
- Dellicour, S., and J. F. Flot. 2015. Delimiting species-poor data sets using single molecular markers: A study of barcode gaps, haplowebs and GMYC. *Syst Biol* 64:900–908.
- De-Silva, D. L., M. Elias, K. Willmott, J. Mallet, and J. J. Day. 2016. Diversification of clearwing butterflies with the rise of the Andes. *J Biogeogr* 43:44–58.

- DeWitt, T. J., and S. M. Scheiner. 2004. Phenotypic variation from single genotypes. Pp. 1–10 in *Phenotypic Plasticity: Functional and conceptual approaches*.
- di Cola, V., O. Broennimann, B. Petitpierre, F. T. Breiner, M. D'Amen, C. Randin, R. Engler, J. Pottier, D. Pio, A. Dubuis, L. Pellissier, R. G. Mateo, W. Hordijk, N. Salamin, and A. Guisan. 2017. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. *Ecography* 40:774–787.
- Dobson, A. A. P., J. P. Rodriguez, W. M. Roberts, and D. S. Wilcove. 2008. Geographic Distribution of Endangered Species in the United States. *Adv Sci* 275:550–553.
- Dowle, E. J., M. Morgan-Richards, and S. A. Trewick. 2014. Morphological differentiation despite gene flow in an endangered grasshopper. *BMC Evol Biol* 14:216.
- Dunn, E. 1926. *The Salamanders of the Family Plethodontidae*. Smith College, Northampton, MA.
- Eaton, D. A. R. 2014. PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30:1844–1849.
- Eck, M. A., L. B. Perry, P. T. Soulé, J. W. Sugg, and D. K. Miller. 2019. Winter climate variability in the southern Appalachian Mountains, 1910–2017. *International Journal of Climatology* 39:206–217.
- Elith, J., M. Kearney, and S. Phillips. 2010. The art of modelling range-shifting species. *Methods Ecol Evol* 1:330–342.
- Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57.
- Epps, C. W., and N. Keyghobadi. 2015. Landscape genetics in a changing world: Disentangling historical and contemporary influences and inferring change. *Mol Ecol* 24:6021–6040.
- Escobar, L. E., H. Qiao, J. Cabello, and A. T. Peterson. 2018. Ecological niche modeling re-examined: A case study with the Darwin's fox. *Ecol Evol* 8:4757–4770.
- Espíndola, A., M. Ruffley, M. L. Smith, B. C. Carstens, D. C. Tank, and J. Sullivan. 2016. Identifying cryptic diversity with predictive phylogeography. *Proceedings of the Royal Society B: Biological Sciences* 283:20161529.
- Excoffier, L., I. Dupanloup, E. Huerta-Sánchez, V. C. Sousa, and M. Foll. 2013. Robust demographic inference from genomic and SNP data. *PLoS Genet* 9:e1003905.
- Excoffier, L., T. Hofer, and M. Foll. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity (Edinb)* 103:285–298.
- Eyring, V., P. M. Cox, G. M. Flato, P. J. Gleckler, G. Abramowitz, P. Caldwell, W. D. Collins, B. K. Gier, A. D. Hall, F. M. Hoffman, G. C. Hurtt, A. Jahn, C. D. Jones, S. A. Klein, J. P. Krasting, L. Kwiatkowski, R. Lorenz, E. Maloney, G. A. Meehl, A. G. Pendergrass, R. Pincus, A. C. Ruane, J. L. Russell, B. M. Sanderson, B. D. Santer, S. C. Sherwood, I. R. Simpson, R. J. Stouffer, and M. S. Williamson. 2019. Taking climate model evaluation to the next level. *Nat Clim Chang* 9:102–110.
- Ezard, T. H. G., and A. Purvis. 2016. Environmental changes define ecological limits to species richness and reveal the mode of macroevolutionary competition. *Ecol Lett* 19:899–906.
- Feldmeier, S., B. R. Schmidt, N. E. Zimmermann, M. Veith, G. F. Ficetola, and S. Lötters. 2020. Shifting aspect or elevation? The climate change response of ectotherms in a complex mountain topography. *Divers Distrib* 26:1483–1495.

- Felsenstein, J. 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* (N Y) 39:783.
- Feng, X., D. S. Park, Y. Liang, R. Pandey, and M. Papeş. 2019. Collinearity in ecological niche modeling: Confusions and challenges. *Ecol Evol* 9:10365–10376.
- Ferchaud, A. L., and M. M. Hansen. 2016. The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: Three-spine sticklebacks in divergent environments. *Mol Ecol* 25:238–259.
- Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* 37:4302–4315.
- Fisher-Reid, M. C., and J. J. Wiens. 2015. Is geographic variation within species related to macroevolutionary patterns between species? *J Evol Biol* 28:1502–1515.
- Fitzpatrick, B. M., J. S. Placyk, Jr, M. L. Niemiller, G. S. Casper, and G. M. Burghardt. 2008. Distinctiveness in the face of gene flow: Hybridization between specialist and generalist gartersnakes. *Mol Ecol* 17:4107–4117.
- Flatley, W. T., C. W. Lafon, H. D. Grissino-Mayer, and L. B. LaForest. 2013. Fire history, related to climate and land use in three southern Appalachian landscapes in the eastern United States. *Ecological Applications* 23:1250–1266. John Wiley & Sons, Ltd.
- Flot, J.-F., A. Couloux, and S. Tillier. 2010. Haplowebs as a graphical tool for delimiting species: a revival of Doyle’s “field for recombination” approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evol Biol* 10:372.
- Fourcade, Y., J. O. Engler, D. Rödder, and J. Secondi. 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. *PLoS One* 9:1–13.
- Francis, R. M. 2017. pophelper: an R package and web app to analyse and visualize population structure. *Mol Ecol Resour* 17:27–32.
- Fraterrigo, J. M., M. G. Turner, S. M. Pearson, and P. Dixon. 2005. Effects of past land use on spatial heterogeneity of soil nutrients in southern Appalachian forests. *Ecol Monogr* 75:215–230. John Wiley & Sons, Ltd.
- French, H. M., and S. W. S. Millar. 2014. Permafrost at the time of the Last Glacial Maximum (LGM) in North America. *Boreas* 43:667–677.
- Frichot, E., F. Mathieu, T. Trouillon, G. Bouchard, and O. François. 2014. Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196:973–983.
- Gabriel, K. R., and R. R. Sokal. 1969. A new statistical approach to geographic variation analysis. *Syst Zool* 18:259–278.
- Galante, P. J., B. Alade, R. Muscarella, S. A. Jansa, S. M. Goodman, and R. P. Anderson. 2017. The challenge of modeling niches and distributions for data-poor species: a comprehensive approach to model complexity. *Ecography* 726–736.
- Gilliam, F. S. 2016. Forest ecosystems of temperate climatic regions: from ancient use to climate change. *New Phytologist* 212:871–887.
- Glor, R. E., and D. Warren. 2011. Testing ecological explanations for biogeographic boundaries. *Evolution* (N Y) 65:673–683.
- Gonzales, E., J. L. Hamrick, and S. M. Chang. 2008. Identification of glacial refugia in south-eastern North America by phylogeographical analyses of a forest understorey plant, *Trillium cuneatum*. *J Biogeogr* 35:844–852.

- Goslee, S. C. 2010. Correlation analysis of dissimilarity matrices. *Plant Ecol* 206:279–286.
- Goslee, S. C., and D. L. Urban. 2007. The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *J Stat Softw* 22:1–19.
- Gragson, T. L., and P. v. Bolstad. 2006. Land use legacies and the future of Southern Appalachia. *Soc Nat Resour* 19:175–190.
- Graham, C. H., S. R. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* (N Y) 58:1781–1793.
- Grant, E. H. C., A. B. Brand, S. F. J. de Wekker, T. R. Lee, and J. E. B. Wofford. 2018. Evidence that climate sets the lower elevation range limit in a high-elevation endemic salamander. *Ecol Evol* 8:7553–7562.
- Guerrero, R. F., and M. W. Hahn. 2017. Speciation as a sieve for ancestral polymorphism. *Mol Ecol* 2017:1–7.
- Guevara, L., B. E. Gerstner, J. M. Kass, and R. P. Anderson. 2018. Toward ecologically realistic predictions of species distributions: A cross-time example from tropical montane cloud forests. *Glob Chang Biol* 24:1511–1522.
- Guisan, A., B. Petitpierre, O. Broennimann, C. Daehler, and C. Kueffer. 2014. Unifying niche shift studies: Insights from biological invasions. *Trends Ecol Evol* 29:260–269. Elsevier Ltd.
- Haddad, N. M., L. A. Brudvig, J. Clobert, K. F. Davies, A. Gonzalez, R. D. Holt, T. E. Lovejoy, J. O. Sexton, M. P. Austin, C. D. Collins, W. M. Cook, E. I. Damschen, R. M. Ewers, B. L. Foster, C. N. Jenkins, A. J. King, W. F. Laurance, D. J. Levey, C. R. Margules, B. A. Melbourne, A. O. Nicholls, J. L. Orrock, D. X. Song, and J. R. Townshend. 2015. Habitat fragmentation and its lasting impact on Earth’s ecosystems. *Sci Adv* 1:1–10.
- Hahn, M. W., and L. Nakhleh. 2016. Irrational exuberance for resolved species trees. *Evolution* (N Y) 70:7–17.
- Hally, M. K., E. M. Rasch, H. R. Mainwaring, and R. C. Bruce. 1986. Cytophotometric evidence of variation in genome size of desmognathine salamanders. *Histochemistry* 85:185–192.
- Halstead, B. J., A. M. Ray, E. Muths, E. H. C. Grant, R. Grasso, M. J. Adams, K. S. Delaney, J. Carlson, and B. R. Hossack. 2022. Looking ahead, guided by the past: The role of U.S. national parks in amphibian research and conservation. *Ecol Indic* 136:108631. Elsevier Ltd.
- Hanken, J. 1983a. High incidence of limb skeletal variants in a peripheral population of the red-backed salamander, *Plethodon cinereus* (Amphibia: Plethodontidae), from Nova Scotia. *Can J Zool* 61:1925–1931.
- Hanken, J. 1983b. Miniaturization and its effects on cranial morphology in plethodontid salamanders, genus *Thorius* (Amphibia: Plethodontidae). I. Osteological variation. *J Morphol* 268:55–75.
- Hanken, J. 1983c. Miniaturization and its effects on cranial morphology in plethodontid salamanders, genus *Thorius* (Amphibia, Plethodontidae): II. The fate of the brain and sense organs and their role in skull morphogenesis and evolution. *J Morphol* 177:255–268.

- Hanks, E. M., and M. B. Hooten. 2013. Circuit theory and model-based inference for landscape connectivity. *J Am Stat Assoc* 108:22–33.
- Hantak, M. M., and S. R. Kuchta. 2018. Predator perception across space and time: Relative camouflage in a colour polymorphic salamander. *Biological Journal of the Linnean Society* 123:21–33.
- Hedin, M., D. Carlson, and F. Coyle. 2015. Sky island diversification meets the multispecies coalescent - Divergence in the spruce-fir moss spider (*Microhexura montivaga*, Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. *Mol Ecol* 24:3467–3484.
- Hedrick, P. W. 2013. Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Mol Ecol* 22:4606–4618.
- Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.
- Hewitt, G. M. 2011. Quaternary phylogeography: the roots of hybrid zones. *Genetica* 139:617–638.
- Hickerson, M. J., B. C. Carstens, J. Cavender-Bares, K. A. Crandall, C. H. Graham, J. B. Johnson, L. Rissler, P. F. Victoriano, and A. D. Yoder. 2009. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Mol Phylogenet Evol* 54:291–301. Elsevier Inc.
- Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
- Hijmans, R. J., S. J. Phillips, J. Leathwick, and J. Elith. 2017. dismo: Species distribution modeling. <http://cran.r-project.org/web/packages/dismo/index.html>.
- Hinderstein, B. 1971. Studies on the salamander genus *Desmognathus*: Variation of lactate dehydrogenase. *Copeia* 1971:636–644.
- Hirzel, A. H., G. le Lay, V. Helfer, C. Randin, and A. Guisan. 2006. Evaluating the ability of habitat suitability models to predict species presences. *Ecol Modell* 199:142–152.
- Ho, S. Y. W. 2014. The changing face of the molecular evolutionary clock. *Trends Ecol Evol* 29:496–503.
- Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh. 2018. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol Biol Evol* 35:518–522.
- Hodkinson, B. P. 2010. A first assessment of lichen diversity for one of North America's 'Biodiversity Hotspots' in the Southern Appalachians of Virginia. *Castanea* 75:126–133.
- Hoffacker, M. Lou, K. K. Cecala, J. R. Ennen, S. M. Mitchell, and J. M. Davenport. 2018. Interspecific interactions are conditional on temperature in an Appalachian stream salamander community. *Oecologia* 188:623–631.
- Hughes, C. E., and G. W. Atchison. 2015. The ubiquity of alpine plant radiations: From the Andes to the Hengduan Mountains. *New Phytologist* 207:275–282.
- Ikeda, D. H., T. L. Max, G. J. Allan, M. K. Lau, S. M. Shuster, and T. G. Whitham. 2017. Genetically informed ecological niche models improve climate change predictions. *Glob Chang Biol* 23:164–176.

- Jackson, N. D. 2005. Phylogenetic history, morphological parallelism, and speciation in a complex of Appalachian salamanders (genus *Desmognathus*). Brigham Young University, Provo, Utah.
- Jaenisch, R., and A. Bird. 2003. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nature Gen. Suppl.* 33:245–254.
- Jamie, G. A., and J. I. Meier. 2020. The persistence of polymorphisms across species radiations. *Trends Ecol Evol* 35:795–808.
- Jarnevich, C. S., M. Talbert, J. Morisette, C. Aldridge, C. S. Brown, S. Kumar, D. Manier, C. Talbert, and T. Holcombe. 2017. Minimizing effects of methodological decisions on interpretation and prediction in species distribution studies: An example with background selection. *Ecol Modell* 363:48–56.
- Jefferson, D. M., M. C. O. Ferrari, A. Mathis, K. A. Hobson, E. R. Britzke, A. L. Crane, A. R. Blaustein, and D. P. Chivers. 2014. Shifty salamanders: Transient trophic polymorphism and cannibalism within natural populations of larval ambystomatid salamanders. *Front Zool* 11:1–11.
- Jombart, T. 2008. Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405.
- Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, J. Baldwin, T. Bloom, D. B. Jaffe, R. Nicol, J. Wilkinson, E. S. Lander, F. Di Palma, K. Lindblad-Toh, and D. M. Kingsley. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.
- Jones, K. S., and D. W. Weisrock. 2018. Genomic data reject the hypothesis of sympatric ecological speciation in a clade of *Desmognathus* salamanders. *Evolution (N Y)* 72:2378–2393.
- Jones, M. T., S. R. Voss, M. B. Ptacek, D. W. Weisrock, and D. W. Tonkyn. 2006. River drainages and phylogeography: An evolutionary significant lineage of shovel-nosed salamander (*Desmognathus marmoratus*) in the southern Appalachians. *Mol Phylogenet Evol* 38:280–287.
- Kadmon, R., O. Farber, and A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications* 14:401–413.
- Kalinowski, S. T. 2010. How to use SNPs and other diagnostic diallelic genetic markers to identify the species composition of multi-species hybrids. *Conserv Genet Resour* 2:63–66.
- Kalinowski, S. T., B. J. Novak, D. P. Drinan, R. Dem Jennings, and N. v. Vu. 2011. Diagnostic single nucleotide polymorphisms for identifying westslope cutthroat trout (*Oncorhynchus clarki lewisi*), Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and rainbow trout (*Oncorhynchus mykiss*). *Mol Ecol Resour* 11:389–393.
- Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermin. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat Methods* 14:587–589.

- Kamvar, Z. N., J. C. Brooks, and N. J. Grünwald. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front Genet* 6:208.
- Kapli, P., Z. Yang, and M. J. Telford. 2020. Phylogenetic tree building in the genomic age. *Nat Rev Genet* 21:428–444.
- Karger, D. N., O. Conrad, J. Böhner, T. Kawohl, H. Kreft, R. W. Soria-Auza, N. E. Zimmermann, H. P. Linder, and M. Kessler. 2017. Climatologies at high resolution for the earth's land surface areas. *Sci Data* 4:1–20.
- Karger, D. N., M. P. Nobis, S. Normand, C. H. Graham, N. E. Zimmermann, and D. N. Karger. 2021. CHELSA-TraCE21k v1.0. Downscaled transient temperature and precipitation data since the last glacial maximum. *Climate of the Past Discussions* 1–27.
- Karger, D. N., D. R. Schmatz, G. Dettling, and N. E. Zimmermann. 2020. High-resolution monthly precipitation and temperature time series from 2006 to 2100. *Scientific Data* 2020 7:1 7:1–10.
- Katoh, K., K. Misawa, K. I. Kuma, and T. Miyata. 2002. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066.
- Keinath, M. C., V. A. Timoshevskiy, N. Y. Timoshevskaya, P. A. Tsonis, S. R. Voss, and J. J. Smith. 2015. Initial characterization of the large genome of the salamander *Ambystoma mexicanum* using shotgun and laser capture chromosome sequencing. *Sci Rep* 5:16413.
- Kiffney, P. M., C. M. Greene, J. E. Hall, and J. R. Davies. 2006. Tributary streams create spatial discontinuities in habitat, biological productivity, and diversity in mainstem rivers. *Canadian Journal of Fisheries and Aquatic Sciences* 63:2518–2530.
- Kimura, M. 1981. Estimation of evolutionary distances between homologous nucleotide sequences. *Genetics* 78:454–458.
- Kjeldsen, S. R., H. W. Raadsma, K. A. Leigh, J. R. Tobey, D. Phalen, A. Krockenberger, W. A. Ellis, E. Hynes, D. P. Higgins, and K. R. Zenger. 2019. Genomic comparisons reveal biogeographic and anthropogenic impacts in the koala (*Phascolarctos cinereus*): a dietary-specialist species distributed across heterogeneous environments. *Heredity (Edinb)* 122:525–544.
- Knowles, L. L., B. C. Carstens, and M. L. L. Keat. 2007. Coupling genetic and ecological-niche models to examine how past population distributions contribute to divergence. *Current Biology* 17:940–946.
- Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour* 15:1179–1191.
- Kozak, K. H., A. A. Larson, R. M. Bonett, and L. J. Harmon. 2005. Phylogenetic analysis of ecomorphological divergence, community structure, and diversification rates in dusky salamanders (Plethodontidae: Desmognathus). *Evolution (N Y)* 59:2000–2016.
- Kozak, K. H., R. W. Mendyk, and J. J. Wiens. 2009. Can parallel diversification occur in sympatry? Repeated patterns of body-size evolution in coexisting clades of North American salamanders. *Evolution (N Y)* 63:1769–1784.

- Kozak, K. H., and J. J. Wiens. 2006. Does niche conservatism promote speciation? A case study in North American salamanders. *Evolution* 60:2604–2621.
- Kratovil, J. D. 2017. Mitochondrial and nuclear patterns of conflict and concordance at the gene, genome, and behavioral scales in *Desmognathus* salamanders. University of Kentucky.
- Kuhman, T. R., S. M. Pearson, and M. G. Turner. 2011. Agricultural land-use history increases non-native plant invasion in a southern Appalachian forest a century after abandonment. *Canadian Journal of Forest Research* 41:920–929.
- Kumar, S., A. J. Filipski, F. U. Battistuzzi, S. L. Kosakovsky Pond, and K. Tamura. 2012. Statistics and truth in phylogenomics. *Mol Biol Evol* 29:457–472.
- La Sorte, F. A., and W. Jetz. 2010. Projected range contractions of montane biodiversity under global warming. *Proceedings of the Royal Society B: Biological Sciences* 277:3401–3410.
- Lamichhaney, S., J. Berglund, M. S. Almén, K. Maqbool, M. Grabherr, A. Martinez-Barrio, M. Promerová, C. J. Rubin, C. Wang, N. Zamani, B. R. Grant, P. R. Grant, M. T. Webster, and L. Andersson. 2015. Evolution of Darwin’s finches and their beaks revealed by genome sequencing. *Nature* 518:371–375.
- Landguth, E. L., S. A. Cushman, M. K. Schwartz, K. S. McKelvey, M. Murphy, and G. Luikart. 2010. Quantifying the lag time to detect barriers in landscape genetics. *Mol Ecol* 19:4179–4191.
- Lanfear, R., J. J. Welch, and L. Bromham. 2010. Watching the clock: Studying variation in rates of molecular evolution between species. *Trends Ecol Evol* 25:495–503.
- Lawson, D. J., L. van Dorp, and D. Falush. 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat Commun* 9:1–11. Springer US.
- Lee-Yaw, J. A., H. M. Kharouba, M. Bontrager, C. Mahony, A. M. Csergo, A. M. E. Noreen, Q. Li, R. Schuster, and A. L. Angert. 2016. A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecol Lett* 19:710–722.
- Leigh, J. W., and D. Bryant. 2015. POPART: Full-feature software for haplotype network construction. *Methods Ecol Evol* 6:1110–1116.
- Leinonen, R., H. Sugawara, and M. Shumway. 2011. The sequence read archive. *Nucleic Acids Res* 39:2010–2012.
- Lemmon, A. R., S. A. Emme, and E. M. Lemmon. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst Biol* 61:727–744.
- Lennox, J. G. 1991. Darwinian thought experiments: A function for just-so stories. Pp. 223–246 in T. Horowitz and G. Massey, eds. *Thought Experiments in Science and Philosophy*. Rowman and Littlefield.
- Lesser, M. R., and J. D. Fridley. 2016. Global change at the landscape level: Relating regional and landscape-scale drivers of historical climate trends in the Southern Appalachians. *International Journal of Climatology* 36:1197–1209.
- Levis, N. A., A. J. Isdaner, and D. W. Pfennig. 2018. Morphological novelty emerges from pre-existing phenotypic plasticity. *Nat Ecol Evol* 2:1289–1297.
- Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 00:1–3.
- Li, X., H. Tian, S. Pan, and C. Lu. 2022. Four-century history of land transformation by humans in the United States: 1630–2020. *Earth Science Data* 1–36.

- Linck, E., and C. J. Battey. 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Mol Ecol Resour* 19:639–647.
- Lindgren, A., G. Hugelius, and P. Kuhry. 2018. Extensive loss of past permafrost carbon but a net accumulation into present-day soils. *Nature* 560:219–222.
- Lindgren, A., G. Hugelius, P. Kuhry, T. R. Christensen, and J. Vandenberghe. 2016. GIS-based maps and area estimates of northern hemisphere permafrost extent during the Last Glacial Maximum. *Permafrost Periglacial Process* 27:6–16.
- Lindtke, D., K. Lucek, V. Soria-Carrasco, R. Villoutreix, T. E. Farkas, R. Riesch, S. R. Dennis, Z. Gompert, and P. Nosil. 2017. Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. *Mol Ecol* 26:6189–6205.
- Liu, C., G. Newell, and M. White. 2016. On the selection of thresholds for predicting species occurrence with presence-only data. *Ecol Evol* 6:337–348.
- Losos, J. B., and L. D. Malher. 2010. Adaptive radiation: The interaction of ecological opportunity, adaptation, and speciation. Pp. 381–420 in M. A. Bell, D. J. Futuyma, W. F. Eanes, and M. A. Levinton, eds. *Evolution Since Darwin: The First 150 Years*. Sinauer Associates, Sunderland, MA.
- Lowe, W. H. 2012. Climate change is linked to long-term decline in a stream salamander. *Biol Conserv* 145:48–53.
- Lowe, W. H., R. P. Kovach, and F. W. Allendorf. 2017. Population genetics and demography unite ecology and evolution. *Trends Ecol Evol* 32:141–152.
- Lozano-Fernandez, J. 2022. A practical guide to design and assess a phylogenomic study. *Genome Biol Evol* 14:1–21.
- Maigret, T. A., J. J. Cox, and D. W. Weisrock. 2020. A spatial genomic approach identifies time lags and historical barriers to gene flow in a rapidly fragmenting Appalachian landscape. *Mol Ecol* 29:673–685.
- Malde, K., B. B. Seliussen, M. Quintela, G. Dahle, F. Besnier, H. J. Skaug, N. Øien, H. K. Solvang, T. Haug, R. Skern-Mauritzen, N. Kanda, L. A. Pastene, I. Jonassen, and K. A. Glover. 2017. Whole genome resequencing reveals diagnostic markers for investigating global migration and hybridization between minke whale species. *BMC Genomics* 18:1–11.
- Mallet, J. 2007. Hybrid speciation. *Nature* 446:279–283.
- Manni, F., E. Guerard, and E. Heyer. 2004. Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Hum Biol* 76:173–190.
- Marques, D. A., K. Lucek, J. I. Meier, S. Mwaiko, C. E. Wagner, L. Excoffier, and O. Seehausen. 2016. Genomics of rapid incipient speciation in sympatric threespine stickleback. *PLoS Genet* 12:1–34.
- Martin, C. H., J. S. Cutler, J. P. Friel, C. Denning Touokong, G. Coop, and P. C. Wainwright. 2015. Complex histories of repeated gene flow in Cameroon crater lake cichlids cast doubt on one of the clearest examples of sympatric speciation. *Evolution (N Y)* 69:1406–1422.
- Martin, C. H., and L. C. Feinstein. 2014. Novel trophic niches drive variable progress towards ecological speciation within an adaptive radiation of pupfishes. *Mol Ecol* 23:1846–1862.

- Martínez-Cabrera, H. I., and P. R. Peres-Neto. 2013. Shifts in climate foster exceptional opportunities for species radiation: The case of South African geraniums. *PLoS One* 8:e83087.
- Martof, B. 1956. Three new subspecies of *Leurognathus marmorata* from the southern Appalachian mountains. *Occas Pap Mus Zool Univ Mich* 1–14.
- Martof, B. S. 1962. Some aspects of the life history and ecology of the salamander *Leurognathus*. *American Midland Naturalist* 67:1–35.
- Mason, N. A., and S. A. Taylor. 2015. Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. *Mol Ecol* 24:3009–3025.
- McCain, C. M., and R. K. Colwell. 2011. Assessing the threat to montane biodiversity from discordant shifts in temperature and precipitation in a changing climate. *Ecol Lett* 14:1236–1245.
- McCarthy, K. P., R. J. Fletcher, C. T. Rota, and R. L. Hutto. 2012. Predicting species distributions from samples collected along roadsides. *Conservation Biology* 26:68–77.
- McLachlan, J. S., J. S. Clark, and P. S. Manos. 2005. Molecular indicators of tree migration capacity under rapid climate change. *Ecology* 86:2088–2098.
- McRae, B. H. 2006. Isolation by resistance. *Evolution (N Y)* 60:1551–1561. Wiley.
- McRae, B. H., and P. Beier. 2007. Circuit theory predicts gene flow in plant and animal populations. *Proc Natl Acad Sci U S A* 104:19885–19890.
- McVay, J. D., A. L. Hipp, and P. S. Manos. 2017. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. *Proceedings of the Royal Society B: Biological Sciences* 284:20170300.
- McVean, G. 2009. A genealogical interpretation of principal components analysis. *PLoS Genet* 5.
- Meirmans, P. G. 2015. Seven common mistakes in population genetics and how to avoid them. *Mol Ecol* 24:3223–3231.
- Meirmans, P. G. 2012. The trouble with isolation by distance. *Mol Ecol* 21:2839–2846.
- Meirmans, P. G., and P. W. Hedrick. 2011. Assessing population structure: F and related measures. *Mol Ecol Resour* 11:5–18.
- Merow, C., M. J. Smith, and J. A. Silander. 2013. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* 36:1058–1069.
- Micheletti, S. J., and A. Storfer. 2017. An approach for identifying cryptic barriers to gene flow that limit species' geographic ranges. *Mol Ecol* 26:490–504.
- Milanovich, J. R., W. E. Peterman, N. P. Nibbelink, and J. C. Maerz. 2010. Projected loss of a salamander diversity hotspot as a consequence of projected global climate change. *PLoS One* 5:e12189.
- Minh, B. Q., M. W. Hahn, and R. Lanfear. 2020a. New methods to calculate concordance factors for phylogenomic datasets. *Mol Biol Evol* 37:2727–2733.
- Minh, B. Q., M. W. Hahn, and R. Lanfear. 2020b. New methods to calculate concordance factors for phylogenomic datasets. *Mol Biol Evol* 37:2727–2733.
- Minh, B. Q., M. A. T. Nguyen, and A. von Haeseler. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* 30:1188–1195.

- Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. Von Haeseler, R. Lanfear, and E. Teeling. 2020c. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 37:1530–1534.
- Mongiardino Koch, N. 2021. Phylogenomic Subsampling and the Search for Phylogenetically Reliable Loci. *Mol Biol Evol* 38:4025–4038.
- Monmonier, M. S. 1973. Maximum-difference barriers: An alternative numerical regionalization method. *Geogr Anal* 5:245–261.
- Morales, A. E., and B. C. Carstens. 2018. Evidence that *Myotis lucifugus* “subspecies” are five nonsister species, despite gene flow. *Syst Biol* 67:756–769.
- Morelli, T. L., A. B. Smith, A. N. Mancini, E. A. Balko, C. Borgerson, R. Dolch, Z. Farris, S. Federman, C. D. Golden, S. M. Holmes, M. Irwin, R. L. Jacobs, S. Johnson, T. King, S. M. Lehman, E. E. Louis, A. Murphy, H. N. T. Randriahaino, H. L. L. Randrianarimanana, J. Ratsimbazafy, O. H. Razafindratsima, and A. L. Baden. 2020. The fate of Madagascar’s rainforest habitat. *Nat Clim Chang* 10:89–96.
- Moskwick, M. 2014. Recent elevational range expansions in plethodontid salamanders (Amphibia: Plethodontidae) in the southern appalachian mountains. *J Biogeogr* 41:1957–1966.
- Murali, G., S. Mallick, and U. Kodandaramaiah. 2021. Background complexity and optimal background matching camouflage. *Behav Ecol Sociobiol* 75. Behavioral Ecology and Sociobiology.
- Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. M. Kass, M. Uriarte, and R. P. Anderson. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol Evol* 5:1198–1205.
- Nadeau, C. P., M. C. Urban, and J. R. Bridle. 2017. Climates past, present, and yet-to-come shape climate change vulnerabilities. *Trends Ecol Evol* 32:786–800.
- Nagaraju, S. K., R. Gudasalamani, N. Barve, J. Ghazoul, G. K. Narayanagowda, and U. S. Ramanan. 2013. Do ecological niche model predictions reflect the adaptive landscape of species?: A test using *Myristica malabarica* Lam., an endemic tree in the Western Ghats, India. *PLoS One* 8:1–13.
- Newbold, T., L. N. Hudson, S. L. L. Hill, S. Contu, I. Lysenko, R. A. Senior, L. Börger, D. J. Bennett, A. Choimes, B. Collen, J. Day, A. de Palma, S. Díaz, S. Echeverria-Londoño, M. J. Edgar, A. Feldman, M. Garon, M. L. K. Harrison, T. Alhusseini, D. J. Ingram, Y. Itescu, J. Kattge, V. Kemp, L. Kirkpatrick, M. Kleyer, D. L. P. Correia, C. D. Martin, S. Meiri, M. Novosolov, Y. Pan, H. R. P. Phillips, D. W. Purves, A. Robinson, J. Simpson, S. L. Tuck, E. Weiher, H. J. White, R. M. Ewers, G. M. MacE, J. P. W. Scharlemann, and A. Purvis. 2015. Global effects of land use on local terrestrial biodiversity. *Nature* 520:45–50.
- Niemiller, M. L., and K. S. Zigler. 2013. Patterns of cave biodiversity and endemism in the Appalachians and Interior Plateau of Tennessee, USA. *PLoS One* 8:e64177.
- Nishimura, K. 2018. An interaction-driven cannibalistic reaction norm. *Ecol Evol* 8:2305–2319.

- Nogués-Bravo, D., M. B. Araújo, M. P. Errea, and J. P. Martínez-Rica. 2007. Exposure of global mountain systems to climate warming during the 21st Century. *Global Environmental Change* 17:420–428.
- Nosil, P., D. J. Funk, and D. Ortiz-Barrientos. 2009. Divergent selection and heterogeneous genomic divergence. *Mol Ecol* 18:375–402.
- Novembre, J. 2016. Pritchard, Stephens, and Donnelly on population structure. *Genetics* 204:391–393.
- Novembre, J., and M. Stephens. 2008. Interpreting principal component analyses of spatial population genetic variation. *Nat Genet* 40:646–649.
- Nowoshilow, S., S. Schloissnig, J.-F. Fei, A. Dahl, A. W. C. Pang, M. Pippel, S. Winkler, A. R. Hastie, G. Young, J. G. Roscito, F. Falcon, D. Knapp, S. Powell, A. Cruz, H. Cao, B. Habermann, M. Hiller, E. M. Tanaka, and E. W. Myers. 2018. The axolotl genome and the evolution of key tissue formation regulators. *Nature* 554:50–55.
- Owens, H. L., and R. Guralnick. 2019. climateStability: An R package to estimate climate stability from time-slice climatologies. *Biodiversity Informatics* 14:8–13.
- Pabijan, M., G. Palomar, B. Antunes, W. Antoń, P. Zieliński, and W. Babik. 2020. Evolutionary principles guiding amphibian conservation. *Evol Appl* 13:857–878.
- Paradis, E. 2018. Analysis of haplotype networks: The randomized minimum spanning tree method. *Methods Ecol Evol* 9:1308–1317.
- Paradis, E. 2010. Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* 26:419–420.
- Parmesan, C. 2006. Ecological and Evolutionary Responses to Recent Climate Change. *Annu Rev Ecol Evol Syst* 37:637–669.
- Pelletier, T. A., C. Crisafulli, S. Wagner, A. J. Zellmer, and B. C. Carstens. 2015. Historical species distribution models predict species limits in western plethodon salamanders. *Syst Biol* 64:909–925.
- Perez, M. F., F. F. Franco, J. R. Bombonato, I. A. S. Bonatelli, G. Khan, M. Romeiro-Brito, A. C. Fegies, P. M. Ribeiro, G. A. R. Silva, and E. M. Moraes. 2018. Assessing population structure in the face of isolation by distance: Are we neglecting the problem? *Divers Distrib* 24:1883–1889.
- Pérez-Escobar, O. A., G. Chomicki, F. L. Condamine, A. P. Karremans, D. Bogarín, N. J. Matzke, D. Silvestro, and A. Antonelli. 2017. Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist* 215:891–905.
- Peterman, W. E., J. A. Crawford, and R. D. Semlitsch. 2008. Productivity and significance of headwater streams: Population structure and biomass of the black-bellied salamander (*Desmognathus quadramaculatus*). *Freshw Biol* 53:347–357.
- Peterson, A. T. 2006. Uses and Requirements of Ecological Niche Models and Related Distributional Models. *Biodiversity Informatics* 3:59–72.
- Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for de novo SNP Discovery and genotyping in model and non-model species. *PLoS One* 7:e37135.
- Petranks, J. W. 2010. *Salamanders of the United States and Canada*. Smithsonian Books, Washington, DC.

- Petranka, J. W., M. P. Brannon, M. E. Hopey, and C. K. Smith. 1994. Effects of timber harvesting on low elevation populations of southern Appalachian salamanders. *For Ecol Manage* 67:135–147.
- Petranka, J. W., M. E. Eldridge, and K. E. Haley. 1993. Effects of timber harvesting on southern Appalachian salamanders. *Conservation Biology* 7:363–370.
- Pfeifer, B., U. Wittelsbürger, S. E. Ramos-Onsins, and M. J. Lercher. 2014. PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Mol Biol Evol* 31:1929–1936.
- Pfennig, D. W., M. A. Wund, E. C. Snell-Rood, T. Cruickshank, C. D. Schlichting, and A. P. Moczek. 2010. Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol Evol* 25:459–467.
- Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259.
- Phillips, S. J., and M. Dudík. 2008. Modeling of species distribution with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175.
- Pickrell, J. K., and J. K. Pritchard. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* 8:e1002967.
- Polo-Cavia, N., and I. Gomez-Mestre. 2017. Pigmentation plasticity enhances crypsis in larval newts: Associated metabolic cost and background choice behaviour. *Sci Rep* 7:1–10.
- Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Mol Biol Evol* 25:1253–1256.
- Priyam, A., B. J. Woodcroft, V. Rai, I. Moghul, A. Munagala, F. Ter, H. Chowdhary, I. Pieniak, L. J. Maynard, M. A. Gibbins, H. K. Moon, A. Davis-Richardson, M. Uludag, N. S. Watson-Haigh, R. Challis, H. Nakamura, E. Favreau, E. A. Gómez, T. Pluskal, G. Leonard, W. Rumpf, and Y. Wurm. 2019. Sequenceserver: A Modern Graphical User Interface for Custom BLAST Databases. *Mol Biol Evol* 36:2922–2924.
- Pyron, R. A., and D. A. Beamer. 2022. Nomenclatural solutions for diagnosing ‘cryptic’ species using molecular and morphological data facilitate a taxonomic revision of the Black-bellied Salamanders (Urodela, *Desmognathus* ‘quadramaculatus’) from the southern Appalachian Mountains. *Bionomina* 27:1–43.
- Pyron, R. A., K. A. O’Connell, E. M. Lemmon, A. R. Lemmon, and D. A. Beamer. 2022. Candidate-species delimitation in *Desmognathus* salamanders reveals gene flow across lineage boundaries, confounding phylogenetic estimation and clarifying hybrid zones. *Ecol Evol* 12:1–38.
- Pyron, R. A., K. A. O’Connell, E. M. Lemmon, A. R. Lemmon, and D. A. Beamer. 2020. Phylogenomic data reveal reticulation and incongruence among mitochondrial candidate species in Dusky Salamanders (*Desmognathus*). *Mol Phylogenet Evol* 146:106751.
- Rabosky, D. L., and I. J. Lovette. 2008. Explosive evolutionary radiations: Decreasing speciation or increasing extinction through time? *Evolution (N Y)* 62:1866–1875.
- Rambaut, A., M. A. Suchard, W. Xie, and A. J. Drummond. 2013. Tracer. <http://tree.bio.ed.ac.uk/software/tracer/>.

- Rannala, B., and Z. Yang. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164:1645–1656.
- Rannala, B., and Z. Yang. 2017. Efficient Bayesian species tree inference under the multi-species coalescent. *Syst Biol* 66:823–842.
- Ravinet, M., A. Westram, K. Johannesson, R. Butlin, C. André, and M. Panova. 2016. Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Mol Ecol* 25:287–305.
- Reddy, S., R. T. Kimball, A. Pandey, P. A. Hosner, M. J. Braun, S. J. Hackett, K. L. Han, J. Harshman, C. J. Huddleston, S. Kingston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, C. C. Witt, T. Yuri, and E. L. Braun. 2017. Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. *Syst Biol* 66:857–879.
- Regos, A., L. Gagne, D. Alcaraz-Segura, J. P. Honrado, and J. Domínguez. 2019. Effects of species traits and environmental predictors on performance and transferability of ecological niche models. *Sci Rep* 9:1–14.
- Reich, D., K. Thangaraj, N. Patterson, A. L. Price, and L. Singh. 2009. Reconstructing Indian population history. *Nature* 461:489–494.
- Remon, J., S. Moulherat, J. H. Cornuau, L. Gendron, M. Richard, M. Baguette, and J. G. Prunier. 2022. Patterns of gene flow across multiple anthropogenic infrastructures: Insights from a multi-species approach. *Landsc Urban Plan* 226.
- Renaut, S., G. L. Owens, and L. H. Rieseberg. 2014. Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Mol Ecol* 23:311–324.
- Rhoads, B. L., and A. N. Sukhodolov. 2001. Field investigation of three-dimensional flow structure at stream confluences: 1. Thermal mixing and time-averaged velocities. *Water Resour Res* 37:2393–2410.
- Riahi, K., D. P. van Vuuren, E. Kriegler, J. Edmonds, B. C. O'Neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, O. Fricko, W. Lutz, A. Popp, J. C. Cuaresma, S. KC, M. Leimbach, L. Jiang, T. Kram, S. Rao, J. Emmerling, K. Ebi, T. Hasegawa, P. Havlik, F. Humpenöder, L. A. da Silva, S. Smith, E. Stehfest, V. Bosetti, J. Eom, D. Gernaat, T. Masui, J. Rogelj, J. Strefler, L. Drouet, V. Krey, G. Luderer, M. Harmsen, K. Takahashi, L. Baumstark, J. C. Doelman, M. Kainuma, Z. Klimont, G. Marangoni, H. Lotze-Campen, M. Obersteiner, A. Tabeau, and M. Tavoni. 2017. The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change* 42:153–168.
- Richards, C. L., B. C. Carstens, and L. Lacey Knowles. 2007. Distribution modelling and statistical phylogeography: An integrative framework for generating and testing alternative biogeographical hypotheses. *J Biogeogr* 34:1833–1845.
- Riddell, E. A., J. P. Odom, J. D. Damm, and M. W. Sears. 2018. Plasticity reveals hidden resistance to extinction under climate change in the global hotspot of salamander diversity. *Sci Adv* 4:1–10.
- Rissler, L. J. 2016. Union of Phylogeography and landscape genetics. *Proc Natl Acad Sci U S A* 113:8079–8086.

- Rissler, L. J., and J. J. Apodaca. 2007. Adding more ecology into species delimitation: Ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). *Syst Biol* 56:924–942.
- Rissler, L. J., and W. H. Smith. 2010. Mapping amphibian contact zones and phylogeographical break hotspots across the United States. *Mol Ecol* 19:5404–5416.
- Rissler, L. J., and D. R. Taylor. 2003. The phylogenetics of Desmognathine salamander populations across the southern Appalachians. *Mol Phylogenet Evol* 27:197–211.
- Robinson, J. T., H. Thorvaldsdóttir, A. M. Wenger, A. Zehir, and J. P. Mesirov. 2017. Variant review with the Integrative Genomics Viewer. *Cancer Res* 77:e31–e34.
- Roesti, M., B. Kueng, D. Moser, and D. Berner. 2015. The genomics of ecological vicariance in threespine stickleback fish. *Nat Commun* 6:1–14.
- Rosa, D., N. Raes, and A. P. Jorge. 2015. Similar but not equivalent: Ecological niche comparison across closely-related Mexican white pines. *Divers Distrib* 21:245–257.
- Ruben, J. A., and A. J. Boucot. 1989. The origin of the lungless salamanders (Amphibia: Plethodontidae). *Am Nat* 134:161–169.
- Ryan, S. F., M. C. Fontaine, J. M. Scriber, M. E. Pfrender, S. T. O’Neil, and J. J. Hellmann. 2017. Patterns of divergence across the geographic and genomic landscape of a butterfly hybrid zone associated with a climatic gradient. *Mol Ecol* 26:4725–4742.
- Samarasin, P., B. J. Shuter, S. I. Wright, and F. H. Rodd. 2017. The problem of estimating recent genetic connectivity in a changing world. *Conservation Biology* 31:126–135.
- Sanderson, B. M., R. Knutti, and P. Caldwell. 2015. A representative democracy to reduce interdependency in a multimodel ensemble. *J Clim* 28:5171–5194.
- Santini, L., A. Benítez-López, L. Maiorano, M. Čengić, and M. A. J. Huijbregts. 2021. Assessing the reliability of species distribution projections in climate change research. *Divers Distrib* 27:1035–1050.
- Sayemuzzaman, M., M. K. Jha, and A. Mekonnen. 2015. Spatio-temporal long-term (1950–2009) temperature trend analysis in North Carolina, United States. *Theor Appl Climatol* 120:159–171.
- Schluter, D. 2001. Ecology and the origin of species. *Trends Ecol Evol* 16:372–380.
- Schmitt, T., and Z. Varga. 2012. Extra-Mediterranean refugia: The rule and not the exception? *Front Zool* 9:1–12.
- Schuler, H., G. R. Hood, S. P. Egan, J. L. Feder, H. Schuler, G. R. Hood, S. P. Egan, and J. L. Feder. 2016. Modes and mechanisms of speciation. 2:60–93.
- Searcy, C. A., and H. B. Shaffer. 2016. Do Ecological Niche Models Accurately Identify Climatic Determinants of Species Ranges? *Am Nat* 187:423–435.
- Seehausen, O. 2004. Hybridization and adaptive radiation. *Trends Ecol Evol* 19:198–207.
- Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, Å. Brännström, A. Brelsford, C. S. Clarkson, F. Eroukhmanoff, J. L. Feder, M. C. Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. A. Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Möst, M. W. Nachman, E. Nonaka, D. J. Rennison, J. Schwarzer, E. T. Watson, A. M. Westram, and A. Widmer. 2014. Genomics and the origin of species. *Nat Rev Genet* 15:176–192.

- Seeholzer, G. F., and R. T. Brumfield. 2018. Isolation by distance, not incipient ecological speciation, explains genetic differentiation in an Andean songbird (Aves: Furnariidae: *Cranioleuca antisiensis*, Line-cheeked Spinetail) despite near threefold body size change across an environmental gradient. *Mol Ecol* 27:279–296.
- Sever, D. M. 1979. Male Secondary Sexual Characters of the *Eurycea bislineata* (Amphibia, Urodela, Plethodontidae) Complex in the Southern Appalachian Mountains. *J Herpetol* 13:245–253.
- Sexton, J. P., S. B. Hangartner, and A. A. Hoffmann. 2014. Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution* (N Y) 68:1–15.
- Shavit, L., D. Penny, M. D. Hendy, and B. R. Holland. 2007. The problem of rooting rapid radiations. *Mol Biol Evol* 24:2400–2411.
- Sillero, N., S. Arenas-Castro, U. Enriquez-Urzelai, C. G. Vale, D. Sousa-Guedes, F. Martínez-Freiría, R. Real, and A. M. Barbosa. 2021. Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. *Ecol Modell* 456:109671.
- Slatkin, M. 1985. Rare alleles as indicators of gene flow. *Evolution* (N Y) 39:53–65.
- Smith, A. B., and M. J. Santos. 2020. Testing the ability of species distribution models to infer variable importance. *Ecography* 43:1801–1813.
- Smith, A. B., M. J. Santos, M. S. Koo, K. M. C. Rowe, K. C. Rowe, J. L. Patton, J. D. Perrine, S. R. Beissinger, and C. Moritz. 2013. Evaluation of species distribution models by resampling of sites surveyed a century ago by Joseph Grinnell. *Ecography* 36:1017–1031.
- Soberón, J. M. 2010. Niche and area of distribution modeling: a population ecology perspective. *Ecography* 33:159–167.
- Sohl, T. L., K. L. Sayler, M. A. Bouchard, R. R. Reker, A. M. Freisz, S. L. Bennett, B. M. Sleeter, R. R. Sleeter, T. Wilson, C. Souldard, M. Knuppe, and T. van Hofwegen. 2018. Conterminous United States Land Cover Projections - 1992 to 2100: U.S. Geological Survey data release.
- Soltis, D. E., A. B. Morris, J. S. McLachlan, P. S. Manos, and P. S. Soltis. 2006. Comparative phylogeography of unglaciated eastern North America. *Mol Ecol* 15:4261–4293.
- Soria-Carrasco, V., Z. Gompert, A. A. Comeault, T. E. Farkas, T. L. Parchman, J. S. Johnston, C. A. Buerkle, J. L. Feder, J. Bast, T. Schwander, S. P. Egan, B. J. Crespi, and P. Nosil. 2014. Stick insect genomes reveal natural selection's role in parallel speciation. *Science* (1979) 344:738–742.
- Spöri, Y., and J. F. Flot. 2020. HaplowebMaker and CoMa: Two web tools to delimit species using haplowebs and conspecificity matrices. *Methods Ecol Evol* 11:1434–1438.
- Steensen, B. M., L. Marelle, Hodnebrog, and G. Myhre. 2022. Future urban heat island influence on precipitation. *Clim Dyn* 58:3393–3403.
- Steyaert, L. T., and R. G. Knox. 2008. Reconstructed historical land cover and biophysical parameters for studies of land-atmosphere interactions within the eastern United States. *Journal of Geophysical Research Atmospheres* 113:1–27.

- Stiven, A. E., and R. C. Bruce. 1988. Ecological genetics of the salamander *Desmognathus quadramaculatus* from disturbed watersheds in the Southern Appalachian Biosphere Reserve Cluster. *Conservation Biology* 2:194–205.
- Storz, J. F. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol* 14:671–688.
- Stroud, J. T., and J. B. Losos. 2016. Ecological Opportunity and Adaptive Radiation. *Annu Rev Ecol Evol Syst* 47:507–532.
- Susko, E., and A. J. Roger. 2021. Long branch attraction biases in phylogenetics. *Syst Biol* 70:838–843.
- Swofford, D. L. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, MA.
- Thesing, B. D., R. D. Noyes, D. E. Starkey, and D. B. Shepard. 2016. Pleistocene climatic fluctuations explain the disjunct distribution and complex phylogeographic structure of the Southern Red-backed Salamander, *Plethodon serratus*. *Evol Ecol* 30:89–104.
- Thibert-Plante, X., and A. P. Hendry. 2011. The consequences of phenotypic plasticity for ecological speciation. *J Evol Biol* 24:326–342.
- Thomson, A. M., C. W. Dick, and S. Dayanandan. 2015. A similar phylogeographical structure among sympatric North American birches (*Betula*) is better explained by introgression than by shared biogeographical history. *J Biogeogr* 42:339–350.
- Thomson, R. C., and J. M. Brown. 2022. On the need for new measures of phylogenomic support. *Syst Biol* 71:917–920.
- Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. *Glob Chang Biol* 10:2020–2027.
- Tilley, S. G., J. Bernardo, L. A. Katz, L. López, J. Devon Roll, R. L. R. L. Eriksen, J. Kratovil, N. K. J. Bittner, and K. A. Crandall. 2013. Failed species, innominate forms, and the vain search for species limits: Cryptic diversity in dusky salamanders (*Desmognathus*) of eastern Tennessee. *Ecol Evol* 3:2547–2567.
- Tilley, S. G., and M. J. Mahoney. 1996. Patterns of genetic differentiation in salamanders of the *Desmognathus ochrophaeus* complex (Amphibia: Plethodontidae). *Herpetological Monographs* 10:1–42.
- Toews, D. P. L., and A. Brelsford. 2012. The biogeography of mitochondrial and nuclear discordance in animals: biogeography of mito-nuclear discordance. *Mol Ecol* 21:3907–3930.
- Toews, D. P. L., S. A. Taylor, R. Vallender, A. Brelsford, B. G. Butcher, P. W. Messer, and I. J. Lovette. 2016. Plumage genes and little else distinguish the genomes of hybridizing warblers. *Current Biology* 26:2313–2318.
- Troscianko, J., O. Nokelainen, J. Skelhorn, and M. Stevens. 2021. Variable crab camouflage patterns defeat search image formation. *Commun Biol* 4:1–9. Springer US.
- Turner, M. G., S. M. Pearson, P. Bolstad, and D. N. Wear. 2003. Effects of land-cover change on spatial pattern of forest communities in the Southern Appalachian Mountains (USA). *Landsc Ecol* 18:449–464.
- Turner, T. L., M. W. Hahn, and S. V. Nuzhdin. 2005. Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol* 3:1572–1578.

- Twyford, A. D., and J. Friedman. 2015. Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. *Evolution* (N Y) 69:1476–1486.
- van Etten, J. 2017. R package gdistance: Distances and routes on geographical grids. *J Stat Softw* 76.
- Vandergast, A. G., B. E. Kus, K. L. Preston, and K. R. Barr. 2019. Distinguishing recent dispersal from historical genetic connectivity in the coastal California gnatcatcher. *Sci Rep* 9:1–12. Springer US.
- VanDerWal, J., L. P. Shoo, C. Graham, and S. E. Williams. 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecol Modell* 220:589–594.
- Varela, S., M. S. Lima-Ribeiro, and L. C. Terribile. 2015. A short guide to the climatic variables of the last glacial maximum for biogeographers. *PLoS One* 10:1–15.
- Vences, M., and D. B. Wake. 2007. Speciation, species boundaries and phylogeography of amphibians. Pp. 2613–2660 in H. Heatwole and M. Tyler, eds. *Amphibian Biology*. Surrey Beatty & Sons, Chipping Norton, Australia.
- Venne, S., and D. J. Currie. 2021. Can habitat suitability estimated from MaxEnt predict colonizations and extinctions? *Divers Distrib* 27:873–886.
- Vodá, R., L. Dapporto, V. Dincă, and R. Vila. 2015. Why do cryptic species tend not to co-occur? A case study on two cryptic pairs of butterflies. *PLoS One* 10:1–18.
- Wake, D. B., M. H. Wake, and C. D. Specht. 2011. Homoplasy: from detecting pattern to determining process in evolution. *Science* (1979) 331:1032–1035.
- Walls, S. C. 2009. The role of climate in the dynamics of a hybrid zone in Appalachian salamanders. *Glob Chang Biol* 15:1903–1910.
- Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. *Nature* 416:389–395.
- Wang, R. J., and M. W. Hahn. 2018. Speciation genes are more likely to have discordant gene trees. *Evol Lett* 2:281–296.
- Warren, D. L. 2012. In defense of “niche modeling.” *Trends Ecol Evol* 27:497–500. Elsevier Ltd.
- Warren, D. L., M. Cardillo, D. F. Rosauer, and D. I. Bolnick. 2014. Mistaking geography for biology: Inferring processes from species distributions. *Trends Ecol Evol* 29:572–580. Elsevier Ltd.
- Warren, D. L., R. E. Glor, and M. Turelli. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* 33:607–611.
- Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution* (N Y) 62:2868–2883.
- Warren, R. J., and M. A. Bradford. 2010. Seasonal Climate Trends, the North Atlantic Oscillation, and Salamander Abundance in the Southern Appalachian Mountain Region. *J Appl Meteorol Climatol* 49:1597–1603.
- Weisrock, D. W., P. M. Hime, S. O. Nunziata, K. Jones, M. O. Murphy, S. Hotaling, and J. D. Kratovil. 2018. Surmounting the large-genome “problem” for genomic data generation in salamanders. P. In press in O. P. Rajora and P. Hohenlohe, eds. *Population Genomics: Wildlife*. Springer International Publishing.

- Wellborn, G. A., and R. B. Langerhans. 2015. Ecological opportunity and the adaptive diversification of lineages. *Ecol Evol* 5:176–195.
- Wiens, J. J. 2012. Why biogeography matters: historical biogeography vs. phylogeography and community phylogenetics for inferring ecological and evolutionary processes. *Front Biogeogr* 4.
- Wiens, J. J., P. T. Chippindale, and D. M. Hillis. 2003. When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. *Syst Biol* 52:501–514.
- Wiens, J. J., C. A. Kuczynski, S. A. Smith, D. G. Mulcahy, J. W. Sites, T. M. Townsend, and T. W. Reeder. 2008. Branch lengths, support, and congruence: Testing the phylogenomic approach with 20 nuclear loci in snakes. *Syst Biol* 57:420–431.
- Wolf, J. B. W., and H. Ellegren. 2017. Making sense of genomic islands of differentiation in light of speciation. *Nat Rev Genet* 18:87–100.
- Wong, A. H. C., I. I. Gottesman, and A. Petronis. 2005. Phenotypic differences in genetically identical organisms: The epigenetic perspective. *Hum Mol Genet* 14:11–18.
- Woodbridge, M., and M. Dovciak. 2022. Logging legacies in a plant biodiversity hotspot: Altered distribution and abundance patterns of the shrub layer in the southern Appalachians. *For Ecol Manage* 516:120245. Elsevier B.V.
- Wooten, J. A., C. D. Camp, and L. J. Rissler. 2010. Genetic diversity in a narrowly endemic, recently described dusky salamander, *Desmognathus folkertsi*, from the southern Appalachian Mountains. *Conservation Genetics* 11:1–20.
- Wooten, J. A., and L. J. Rissler. 2011. Ecological associations and genetic divergence in Black-bellied salamanders (*Desmognathus quadramaculatus*) of the southern Appalachian Mountains. *Acta Herpetol* 6:175–208.
- Wu, C. I. 2001. The genic view of the process of speciation. *J Evol Biol* 14:851–865.
- Yang, Z., and B. Rannala. 2010. Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci U S A* 107:9264–9269.
- Yi, X., and E. K. Latch. 2022. Nonrandom missing data can bias Principal Component Analysis inference of population genetic structure. *Mol Ecol Resour* 22:602–611.
- Zhang, D., G. Q. Hao, X. Y. Guo, Q. J. Hu, and J. Q. Liu. 2019. Genomic insight into “sky island” species diversification in a mountainous biodiversity hotspot. *J Syst Evol* 57:633–645.
- Zhang, Y., C. Chen, L. Li, C. Zhao, W. Chen, and Y. Huang. 2014. Insights from ecological niche modeling on the taxonomic distinction and niche differentiation between the blackspotted and red-spotted tokay geckoes (*Gekko gecko*). *Ecol Evol* 4:3383–3394.
- Zhao, S., Y. Guo, Q. Sheng, and Y. Shyr. 2014. Heatmap3: An improved heatmap package with more powerful and convenient features. *BMC Bioinformatics* 15:15–16.
- Zou, F., S. Lee, M. R. Knowles, and F. A. Wright. 2010. Quantification of population structure using correlated SNPs by shrinkage principal components. *Hum Hered* 70:9–22.

VITA

Kara S. Jones

EDUCATION

2015. M.S. in Environmental Science & Policy, George Mason University
2013. B.S. in Biology, Summa Cum Laude, George Mason University
2011. A.S. in Science, Northern Virginia Community College

TEACHING AND RESEARCH EXPERIENCE

2015-2023. Graduate Teaching Assistant and Research Assistant, University of Kentucky
2013-2015. Graduate Teaching Assistant, George Mason University
2012-2013. Field Interpreter, Potomac Environmental Research & Education Center

PUBLICATIONS

Jones, KS, and DW Weisrock. 2018. Genomic data reveal lineage divergence follows geography, not phenotype, in a clade of sympatric *Desmognathus* salamanders. *Evolution*. 72:2378-2393.

Weisrock, DW, Hime, PM, Nunziata, SO, **Jones, K**, Murphy, MO, Hotaling, S, and JD Kratovil. 2018. Surmounting the large-genome "problem" for genomic data generation in salamanders In *Population Genomics: Wildlife*. Eds Rajora, OP, and P Hohenlohe. Springer Int'l Publishing.

Murphy, MO, **Jones, KS**, Price, SJ, and DW Weisrock. 2018. A genomic assessment of population structure in an aquatic salamander identifies the roles of spatial scale, barriers, and river architecture. *Freshwater Biol.* 63: 407-419.

Jones, KS and TA Tupper. 2017. Aspects of Eastern Wormsnake (*Carphophis amoenus*) microhabitats at two natural areas in Fairfax County, Virginia, and Anne Arundel County, Maryland. *Catesbeiana* 37: 99-108.

Marsh, DM, Cosentino, BJ, **Jones, KS**, et al. 2017. Effects of roads and land use on pond-breeding amphibians: patterns across spatial scales and regions in the Eastern and Central United States. *Divers. Distributions* 23:158-170.

Jones, KS, and TA Tupper. 2015. Fowler's Toad (*Anaxyrus fowleri*) occupancy in the southern mid-Atlantic, USA. *Amphibian & Reptile Conservation* 9: 24-33 (e97).

Tupper, TA, Bozarth, CA, **Jones, KS**, and RP Cook. 2014. Detection of *Batrachochytrium dendrobatidis* in the Eastern Spadefoot, *Scaphiopus holbrookii*, at Cape Cod National Seashore, Barnstable County, Massachusetts, USA. *Herpetol. Rev.* 45:445-447.

Cosentino, BJ, Marsh, DM, **Jones, KS**, et al. 2014. Citizen science reveals widespread negative effects of roads on amphibian distributions. *Biol. Conserv.* 180:31-38.

GRANTS AND AWARDS

2023. Certificate for Outstanding Teaching, College of Arts & Sciences
2019. Appalachian Highlands Science Learning Center Research Program Grant
2018. Morgan Graduate Fellowship
2017. Kentucky Academy of Science Marcia Athey Grant
2016. Society of Systemic Biologists Graduate Research Award

2016. University of Kentucky Ribble Mini-Grant
2015. Kentucky Opportunity Fellowship
2015. Bruce Family Scholarship in Herpetology
2015. Wimberly C. Royster Graduate Excellence Award of Special Distinction*
2015. Gertrude Flora Ribble Research Fellowship
2014. Patriot Green Fund
2013. Undergraduate Research Scholar Grant, George Mason University

* 3 year award, renewed 2016 and 2017