

University of Mississippi

eGrove

Electronic Theses and Dissertations

Graduate School

1-1-2019

Application and refinement of molecular ecology techniques for amphibian conservation

Stephanie Marie Burgess

Follow this and additional works at: <https://egrove.olemiss.edu/etd>



Part of the [Biology Commons](#), [Genetics Commons](#), and the [Natural Resources and Conservation Commons](#)

Recommended Citation

Burgess, Stephanie Marie, "Application and refinement of molecular ecology techniques for amphibian conservation" (2019). *Electronic Theses and Dissertations*. 1956.

<https://egrove.olemiss.edu/etd/1956>

This Dissertation is brought to you for free and open access by the Graduate School at eGrove. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

APPLICATION AND REFINEMENT OF MOLECULAR ECOLOGY TECHNIQUES
FOR AMPHIBIAN CONSERVATION

A Dissertation
presented in partial fulfillment of requirements
for the degree of Doctor of Philosophy
in the Department of Biology
The University of Mississippi

by

STEPHANIE M. BURGESS, MS

AUGUST 2019

Copyright Stephanie M. Burgess 2019

ALL RIGHTS RESERVED

ABSTRACT

Wildlife conservation has become increasingly difficult due to habitat loss, habitat fragmentation, and land use change. Thus, conservationists have embraced advances in molecular ecology, such as landscape genetics and microbial bioinformatics, that employ genetic techniques to further understand the relationship between individuals and their environment. In landscape genetics, model inferences can be used to identify features that facilitate or resist gene flow, providing a framework for anticipating the impacts of land use changes on a species' ability to disperse. However, the factors that affect the transferability of landscape genetics inferences are poorly understood, and little is known about the effect of sampling density and study area size on landscape genetics inferences. To address these understudied factors, I performed a series of landscape genetics analyses using populations of the Mississippi slimy salamander (*Plethodon mississippi*) in Mississippi and Alabama. Regional replication revealed the importance of habitat configuration on the relationship between land use and gene flow among salamander populations, and the transferability of landscape genetics inferences to neighboring areas. Analysis of hierarchically nested datasets of different sampling densities and study area sizes identified differences due to study area size, however no clear effect was seen as a result of different sampling densities. Conservation practitioners can also use microbial ecology to better understand the relationship between wildlife species and their environment. The mutualistic relationship between amphibians and their cutaneous microbial community can strengthen the amphibian's ability to fight fungal pathogens. However, in order to inform

management strategies such as probiotic inoculation, researchers must first understand the method in which amphibian cutaneous microbiomes are shaped. I compared salamander relatedness, salamander cutaneous microbiomes, and the microbiomes of salamanders' immediate soil environment, which revealed no relationship between kinship and similarity of skin microbiomes. Further, comparison of skin and soil microbiomes provided evidence that the presence of antifungal taxa in a salamander's environment does not guarantee incorporation of the taxa into salamander cutaneous microbiomes. The results of this research fill knowledge gaps within the fields of landscape genetics and amphibian cutaneous microbial ecology and provide a greater understanding of the relationship between *P. mississippi* and its environment.

DEDICATION

This dissertation is dedicated to my family, who have endured countless trips to the field and weekends spent writing. Tayler, Asher, and Riley, thank you for supporting me through it all.

I also dedicate this work to all the young girls whose drive and determination are written off as “overzealous” and “emotional” - may your success outshine the doubters and stand as a beacon for those who will follow.

ACKNOWLEDGEMENTS

I thank my advisor, Dr. Ryan Garrick, for his support and counsel throughout this process. I also thank my committee, Dr. Greg Easson, Dr. Brice Noonan, and Dr. William Resetarits, and Dr. Rebecca Symula. I also thank Dr. Colin Jackson for his support throughout my microbial analysis. I thank the Birmingham Audubon Society and the University of Mississippi Graduate Student Council, who provided funding for this research, and the US Forest Service for permitting sampling within Holly Springs and Bankhead National Forests.

I thank my fellow graduate students in the Department of Biology, who have always been happy to discuss new ideas and hear frustrations about statistical software. I also thank Nicole Lewis, who was a joy to work for and also frequently a sympathetic ear. Finally, it is with deepest gratitude that I thank my family: my daughters Asher and Riley, and my supportive and dedicated husband Tayler, without whom none of this would be possible.

TABLE OF CONTENTS

ABSTRACT.....	ii
DEDICATION.....	iv
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: REGIONAL REPLICATION OF LANDSCAPE GENETICS ANALYSES OF THE MISSISSIPPI SLIMY SALAMANDER (<i>PLETHODON MISSISSIPPI</i>)	5
CHAPTER 3: THE EFFECT OF GEOGRAPHIC SAMPLING SCALE ON LANDSCAPE GENETICS INFERENCES FOR THE MISSISSIPPI SLIMY SALAMANDER (<i>PLETHODON MISSISSIPPI</i>)	32
CHAPTER 4: ANALYSIS OF CUTANEOUS AND LOCAL SOIL MICROBIOMES IN A TERRESTRIAL SALAMANDER: THE ROLE OF SIBLINGS VERSUS SOIL	66
CHAPTER 5: CONCLUSION.....	84
REFERENCES.....	87
APPENDIX.....	110
VITA.....	118

LIST OF TABLES

Table 1. Maximum likelihood population effects models and AICc scores.....	16
Table 2. Comparison of landscape metrics for Holly Springs National Forest and Bankhead National Forest.....	21
Table 3. Rank and model coefficients of landscape variables in most supported Maximum likelihood population effects model for each forest region.....	21
Table 4. Comparison of scale, transformation, and sign of effect for Holly Springs National Forest and Bankhead National Forest.....	22
Table 5. AICc scores for each multivariate Maximum Likelihood Population Effects model for each set of analyses.....	50
Table 6. Rank of effect of landscape variables in best-fit maximum likelihood population effects models for each analysis.....	55
Table 7. Comparison of the final optimized transformation, scale, and sign of effect for each land use class in all three datasets.....	55

LIST OF FIGURES

Figure 1. Depiction of landscape classification and optimization for scale and transformation...	18
Figure 2. Comparison of the facilitation or resistance to gene flow created by land use types in Holly Springs National Forest versus Bankhead National Forest.....	23
Figure 3. Sampling locations of <i>P. mississippi</i> within Holly Springs National Forest (HSNF), Mississippi USA.....	39
Figure 4. Different combinations of sampling density and study area size, and hypothetical outcomes relating to similarity of landscape genetics inferences among datasets.....	40
Figure 5. Transformations of original resistance values.....	49
Figure 6. Distribution of the pine land use class in Holly Springs National Forest.....	51
Figure 7. Nested regressions of geographic distance against genetic distance.....	52
Figure 8. Semivariogram created using pairwise genetic distances and geographic Euclidean distances using 52 distance classes with a distance interval of 1.5 km.....	53
Figure 9. Relative abundance by sample of a) common phyla and b) subphyla within Proteobacteria.....	76
Figure 10. Principal Coordinates Analysis (PCoA) of pairwise UNIFRAC distances between all microbiome samples.....	77
Figure 11. Bacterial families with significant differential abundance between paired samples of salamander cutaneous microbiomes and the microbiomes of their immediate environment.....	78
Figure 12. UniFrac distances by kinship category.....	80

CHAPTER 1

INTRODUCTION

As a species, humans have created an impact on our planet that is so substantial many have proposed a new epoch, the Anthropocene, that describes a period of time in which human activity outweighs the forces of nature (Steffen et al. 2007). For wildlife species, this epoch will be characterized by continued habitat loss, habitat fragmentation, and land use change. In light of these challenges, the future of modern conservation has been hotly debated, with advocates for both human centered, anthropocentric approaches and biodiversity centered approaches (Miller et al. 2013; Corlett 2015). No matter their overarching philosophy, conservation practitioners are faced with tough decisions exacerbated by limited funding and conflicting stakeholders.

Although management is increasingly difficult, innovative advancements in molecular ecology offer new techniques to incorporate scientific investigation into management strategies.

However, a gap exists between the overarching, hypothesis-driven questions typical to primary research and the species-specific, local projects that conservation practitioners seek to inform (Braunisch et al. 2012). The studies within this dissertation are designed to address some of the issues that conservation practitioners face when incorporating two techniques of molecular ecology, landscape genetics and microbial bioinformatics, into their decision-making process.

Landscape genetics is an interdisciplinary field wherein researchers attempt to determine the effects of different landscape features on dispersal and gene flow using landscape ecology

and population genetic techniques (Manel et al. 2003). The inferences made in landscape genetics models can be used to anticipate the impacts of current and future land use on a focal species' long-term viability (Sork and Waits 2010). For example, land managers can identify habitat that is effectively isolated due to landscape features that cause decreased dispersal and gene flow and create spatially explicit corridors through these areas (Braunish et al. 2010).

Landscape genetics methods have the ability to generate information faster than conventional scientific methods such as capture-mark-recapture (Berry et al. 2004), however they analyses still require a significant investment of both time and money. State and federal conservation practitioners are often called upon to justify the allocation of limited research funds by demonstrating the overall value and broad applicability proposed projects, including the applicability of research findings to nearby areas within a species' range. Thus, there is a continuing need for empirical landscape genetics research that demonstrates the transferability (i.e. applicability to neighboring areas) of landscape genetics inferences (Short Bull et al. 2011). Furthermore, funds are typically preferentially allocated to research that includes well-defined methodology supported by previous studies as opposed to research whose goal is to develop novel methods. As such, there is also a need for hypothesis-driven landscape genetics studies that empirically test the methodology of the field.

To address methods of transferability, a focal species that is common across a large study region and susceptible to land use change and environmental heterogeneity is essential. The Mississippi slimy salamander, *Plethodon mississippi* (Highton 1989), is a terrestrial salamander that spends the majority of its life under downed woody debris, in caves, or in leaf litter on the

forest floor. *P. mississippi* are commonly found in the bottomland hardwood forests, swamp forests, and wet pine-woods of Mississippi, Alabama, and western Tennessee (Petranka, 1998). As a directly developing species, *P. mississippi* do not need to disperse to aquatic environments for reproduction, and move very little over their lifetimes (Wells and Wells 1976). Because of these life history traits, this low-mobility salamander is a fitting focal species for comparative landscape genetics analyses.

Like many wildlife species in the Anthropocene, salamanders like *P. mississippi* are not only threatened by habitat fragmentation and land use change, they are also at risk from invasive pathogens (Collins and Storfer 2003). Management of disease spread across wildlife populations requires a multi-faceted approach from conservation practitioners that includes both empirical, experimental research and careful response planning (Langwig et al. 2015). Part of this planning must include projections of potential disease spread, which can be informed by an understanding of the microbial communities in and around wildlife species through microbial bioinformatics (Bahrndorff et al. 2016). Mitigation and containment strategies can also be informed by knowledge of host microbiomes. For instance, a promising management strategy for the containment of *Batrachochytrium salamandrivorans* (*Bsal*), a potentially lethal salamander fungus, is the manipulation of the unique relationship between salamanders and the microbial communities of their skin through the introduction of bacteria that exhibit antifungal properties via probiotic mixtures (Becker and Harris 2010). However, in order to understand the probability of uptake of these beneficial microbes, conservation practitioners must first develop an understanding of how salamander cutaneous microbiomes are shaped.

In this collection of research, Chapters 2 and 3 have been designed to fill knowledge gaps in the application of landscape genetics, specifically addressing the transferability (Chapter 2) of model inferences, and the effect of sampling density and study area size (Chapter 3) on model inferences. Chapter 4 explores the possible factors that influence salamander cutaneous microbiome species composition by using relatedness and kinship data, as well as paired comparison of salamander cutaneous microbiomes and the microbiomes of their immediate environment.

CHAPTER 2
REGIONAL REPLICATION OF LANDSCAPE GENETICS ANALYSES OF THE
MISSISSIPPI SLIMY SALAMANDER (*PLETHODON MISSISSIPPI*)

Abstract

Landscape genetics inferences can be used to identify features that facilitate or resist gene flow, providing a framework for anticipating the impacts of land use changes on a species' ability to disperse. To use this framework for management, it is necessary to understand how inferences derived from one region are applicable to other regions within a species' range. We investigated whether the landscape variables assessed in landscape genetics analyses of *Plethodon mississippi* in two different study regions showed the same order of importance, had the same direction and scale of effect, and/or exhibited the same functional relationship to gene flow. In forests in Mississippi and Alabama, USA, we tested individual-based genetic distances derived from microsatellite genotypes against five landscape variables that were optimized for both scale and transformation using maximum likelihood population effects modeling. Of the five landscape variables, agriculture and wetlands ranked at the top of both forests' best-fit models. Whereas agriculture consistently caused resistance, and pine consistently facilitated gene flow across the two forest regions, we found region-specific differences in effects of wetlands, hardwoods, and manmade structures on *P. mississippi* gene flow. Configuration of the latter landscape variables differed between forest regions. Our results underscore the value of

metareplication in revealing which components of landscape genetics models may be consistent across different portions of a species' range, and those that have context-dependent impacts on gene flow. We also highlight the need to consider habitat configuration when interpreting landscape genetics inferences.

Introduction

All species have areas of preferred habitat interspersed with areas of sub-optimal or unsuitable habitat within their range (i.e., a matrix; Fahrig and Merriam 1985). In order to maintain demographic and genetic connectivity among local populations that reside within different habitat patches, individuals must be able to traverse the intervening matrix. However, such areas are increasingly heterogeneous and volatile due to anthropogenic influences. Modifications of natural areas are occurring at an accelerated rate due to the direct effects of a growing human population and associated expansion of urban areas, as well as indirect effects such as alteration of natural disturbance regimes, introduction of exotic species, and climate change (Vitousek et al. 1997; Oswald et al. 2015; Parisien et al. 2016). As a result, areas that were previously comprised mostly of suitable habitat areas have become increasingly "hostile" to free movement of individuals. This change in the permeability of the habitat matrix can lead to long-term isolation among locally small populations and random loss of genetic diversity due to the predominance of drift over selection. As inbreeding becomes unavoidable in small isolated populations, this can give rise to inbreeding depression. In turn, these negative effects on individual fitness and reproductive output further diminish population size and growth rate (Allendorf et al. 2013). Indeed, these population-level changes can interact with other

threatening processes (e.g., rapid changes in the abiotic environment, or emergence of wildlife disease) leading to local extinction (Gilpin and Soulé 1986), and by extension, an overall reduction in a species' long-term viability (Sork and Waits 2010).

Knowledge about the relationship between organisms and their environments is a cornerstone of natural resource management. Wildlife conservation must consider the consequences of population isolation in the design of protected area networks and corridors, and this requires an understanding of the effect of specific landscape features on dispersal of individuals, and gene flow among populations. For decades techniques such as capture-mark-recapture and radio telemetry have been used to gain such insights (e.g., Ovaska 1988; Riecken and Raths 1996). These methods are valuable, but have notable limitations. For example, capture-mark-recapture studies are time and labor intensive, and data points are acquired only from individuals that are re-encountered (Berry et al. 2004). Furthermore, the probability of recapturing marked individuals that have dispersed large distances is very low, creating an observation bias toward detection of short-distance dispersal events (Koenig et al. 1996). Similarly, radio telemetry and passive integrative transponder tagging are also time and labor intensive, and involve expensive equipment such that data are typically obtained from relatively few individuals (Hebblewhite and Haydon 2010; Connette and Semlitsch 2015). While these methods can provide high-resolution information on fine-scale individual movement, given that all data are usually acquired from a single cohort of individuals, capture-mark-recapture and radio telemetry provide only a short temporal snapshot. Accordingly, inferences may be influenced by abnormal environmental conditions, and could be unrepresentative (Bailey et al. 2004).

In principal, a robust understanding of how individuals perceive and move through a habitat matrix would be drawn from a large number of individuals sampled over a range of spatial and temporal scales, with at least two tiers of temporal insights: those reflecting very recent dispersal events (i.e., within the past generation or two), and those based on the accumulated effects of many generations of repeated dispersal and gene flow. Molecular approaches have been used for these purposes, using individual-based comparisons of multilocus genotypes to determine recent dispersal, and population-based allele frequencies to detect the effects of repeated dispersal over time (e.g., Sunnucks 2000; Epps et al. 2013a,b). When employing a landscape genetics approach, molecular data are used to generate genetic distances between individuals or populations, which are then compared to corresponding distances based on the permeability of intervening heterogeneous habitats (Manel et al. 2003). For example, in an early landscape genetics study of gene flow among European roe deer in a fragmented landscape, Coulon et al. (2004) considered two alternative measures of spatial distances: straight line distances versus the path that maximized use of wooded corridors (resistance distance). Those authors found that compared to simple isolation-by-distance, the latter ecologically informed "resistance distance" provided a significantly better fit to inter-individual genetic distances based on microsatellite data, showing that roe deer dispersal is strongly tied to wooded areas.

Today, landscape genetics studies have become more analytically advanced, but the same basic principles apply: the hypothesized resistance to dispersal caused by landscape variables such as land cover, topography, or various bioclimatic measures (i.e., potential predictor variables) is tested against empirically derived genetic distances (i.e., the response variable) in an

effort to determine which landscape features most strongly resist (or, conversely, facilitate) gene flow. Findings from these investigations can be used to plan for the impacts of recent and future land use changes upon a species' ability to disperse, thus providing spatially explicit guidance for conservation management (e.g., Cleary et al. 2017).

There are a variety of spatial data types available in landscape genetics, but of these, land cover classifications, presence or absence of roads, and topographic data are among the most commonly used (Zeller et al. 2012). The decision to include a given landscape variable, and associated choices regarding its hypothesized resistance to gene flow, is typically informed by expert opinion and literature reviews (Beier et al. 2008). While these approaches have value, they may nonetheless overlook relationships that are counterintuitive given the current understanding of organism's natural history (e.g., Peterman et al. 2014). Some of the potential bias associated with relying on *a priori* assumptions to define resistance weightings (i.e., the presumed permeability) of different types of landscape features can be avoided by reassessing the contribution of each landscape variable at multiple geographic scales, and in multiple functional forms. An additional source of potential bias relates to idiosyncrasies associated with the chosen study region. Indeed, understanding the transferability (i.e., applicability to other areas) of landscape genetics models is critical to their use in conservation (Keller et al. 2014), and as such, metareplication is a potentially powerful approach for distinguishing between site-specific versus species-specific processes.

The geographic scale at which individuals of a species perceive habitat quality can be variable and difficult to ascertain (Mayor et al. 2009). For example, a large scale may mean that

a road located 500 m away would nonetheless impede dispersal. Conversely, the same species may interact with pine ridges on a much smaller scale, meaning pine even 100 m away does not exert an effect. In the past, expert opinion, literature review, or habitat suitability assessments have been used to set these scales, often with the same scale uniformly applied to all landscape variables under consideration (McGarigal et al. 2016). However, as noted by Galpern et al. (2012) and Zeller et al. (2017), a more suitable approach would be to consider several alternative geographic scales for each landscape variable in order to determine the appropriate fit. Another approach to model optimization focuses on the functional relationship between a landscape variable and its level of resistance. The function (i.e., transformation) of resistance, like geographic scale, has often been assigned on the basis of expert opinion or literature review (Beier et al. 2008). To date, the most typical functional relationship has been negative and linear. However several studies, including those with genetic response variables (Cushman et al. 2006; Zeller et al. 2017) and with physical animal tracking (Trainor et al. 2013; Keeley et al. 2016), have found support for non-linear functional relationships between landscape variables and resistance. For example, Cushman et al. (2006) modeled a series of Gaussian relationships between elevation and resistance to gene flow in black bears to determine the elevation at which resistance to gene flow among bear populations was the lowest.

Replicated empirical analyses, or metareplications, have the ability to determine how transferable landscape genetics models are across a species' range, and to provide insights into the relationship between model optimization and transferability. Successful metareplication design requires that a species is distributed across a region large enough to have at least two replicate study areas. While these must be similar enough to contain the same study species, it is

important that they not be identical to each other, thereby allowing the researcher to draw conclusions about how the study species responds to landscape variables generally (Johnson 2002).

The present study focused on a species distributed throughout eastern Mississippi and western Alabama, the Mississippi slimy salamander (*Plethodon mississippi* Highton 1989). Plethodontid salamanders represent low-mobility ecologically specialized taxa that have several life history traits that make them well-suited for landscape genetic studies. These amphibians inhabit cool, moist environments (Petranka 1998). They also exhibit direct development, meaning their offspring do not need an aquatic environment to metamorphose into the adult form (Petranka 1998). Without the need to disperse to aquatic environments for reproduction, it is hypothesized they disperse very little over their lifetimes, which may cause genetic differentiation among populations over a relatively small geographic area. Furthermore, due to *P. mississippi*'s short generation time (females and males reach sexual maturity in two years and three years respectively; Highton 1962), the effect on dispersal by changes in the landscape may be detected over relatively short times scales.

The geographic range of *P. mississippi* spans Holly Springs National Forest (HSNF) in northern Mississippi, and Bankhead National Forest (BNF) approximately 190 km to the east in northern Alabama. These two forest regions encompass similar land use types, with both containing bottomland hardwood forests, forested wetlands, upland pine and silviculture, agricultural fields and pastures, and manmade structures such as roads, buildings, and parking lots. While composition of these forest regions is similar and both are managed by the U.S.

Forest Service, they differ in that only BNF includes a large protected Wilderness area (over 25,000 acres). Also, whereas BNF contains roughly 3,500 acres of old growth, HSNF has none (U.S. Department of Agriculture Forest Service 2004, 2012).

In this study, we conducted separate landscape genetic analyses of *P. mississippi* in HSNF and BNF to understand the extent to which inferences drawn from one location are transferable to the other, and to examine the effect of optimization on transferability. We asked if the landscape variables that were assessed (1) show the same rank or order of importance, (2) have the same direction of effect (i.e. facilitate versus resist gene flow), (3) have the same scale of effect, and (4) exhibit the same functional relationship.

Methods

Tail tip tissue was collected from 113 *P. mississippi* individuals at 19 locations in HSNF in northern Mississippi, and 110 individuals at 20 locations in BNF in northern Alabama. Sampling locations were chosen to span the entirety of each of the two forest regions, and spaced approximately eight km apart. At least five individuals were sampled at each location. Average distance between individuals within sampling locations was 122 m. Because *P. mississippi* is a completely terrestrial species that is likely continuously distributed, population units cannot be readily delimited *a priori*. Accordingly, we calculated individual-based genetic distance (Shirk and Cushman 2014).

Genetic Analysis

Genomic DNA was extracted from tail tips using a DNeasy Blood and Tissue kit (Qiagen, Valencia CA, USA) following the manufacturer's recommendations. Individuals were genotyped using eight microsatellite loci described by Spatola et al. (2013; see Appendix for PCR amplification conditions, and allele-calling approaches). At each of three locations in HSNF and one location in BNF we collected 9-11 individuals. These four sample sets were tentatively assumed to each represent panmictic groups for the purpose of testing for null alleles, Hardy Weinberg Equilibrium, and linkage disequilibrium, using Genepop v 4.2 (Raymond and Rousset 1995). Based on the full genetic dataset, the R (R Core Team 2019) package “PopGenReport” (Adamack and Gruber 2014) was used to quantify percent missing data, number of alleles per locus, and mean allelic richness in each forest region. An examination of overall population structure within each forest region was performed via genotypic clustering using STRUCTURE v. 2.3.4 (Prichard et al. 2001). Briefly, we examined K values from 1–5 (3 replicates per K), using the correlated allele frequencies and admixture ancestry models (with alpha and lambda inferred separately for each cluster), with a burn-in of 1×10^5 MCMC iterations, and run length of 1×10^6 iterations. The best fit value of K was identified via comparison of the mean log likelihood of each value of K, and calculation of delta K following Evanno et al. (2005) in STRUCTURE HARVESTER (Earl and vonHoldt 2012). Also within each forest region, we used GenAIEx v. 6.503 (Peakall and Smouse 2012) to test for spatial autocorrelation using, 999 permutations, 999 bootstrap replicates, and tests for heterogeneity. For these analyses, a distance class (i.e., bin size) size of 3 km was chosen to encompass the smallest distances between sampling locations, which were greater than 8 km.

To determine pair-wise individual-based genetic distances within each forest region, we conducted a principal components analysis and calculated Euclidean distance between the first 64 axes of the ordination using the “ade4” package (Dray and Dufour 2007) in R. This method has been shown to perform better than others when genetic structure and sample sizes are low (Shirk et al. 2017). Pairwise genetic distances among individuals from the same sampling location were removed from further analyses to avoid skewing landscape genetics models.

Landscape Analysis

To test the hypothesis that land use type would influence gene flow, we classified spatial data into six distinct land use classes using multi-spectral raster files from the USGS Landsat 8 satellite (see Figure 1, left panel). Through the supervised classification feature in ERDAS Imagine 2014 (Hexagon Geospatial, Norcross GA, USA), each pixel in the 30 x 30 m multispectral image was classified as either agricultural, hardwood, manmade (e.g., paved surfaces and buildings), pine, wetland, and water body land uses using training areas developed using high-resolution imagery and previous knowledge of the study area and a maximum likelihood algorithm. We overlaid wetland, water flowline, and road shapefiles onto the classified image to ensure forested wetlands, small water features, and small roads were included in the classification. The overlay was created using raster calculator in ArcGIS 10.2.2 (ESRI 2011). We used the final maps to calculate the amount of habitat, patch density, correlation length, clumpiness, patch cohesion, and an aggregation index for each landscape variable using the software FRAGSTATS v 4.2 (McGarigal et al. 2012). We then conducted a series of

univariate moving window analyses on the classified images using five separate kernel sizes (100, 250, 500, 750, and 1000 m) for each land use class with the PLAND calculation in FRAGSTATS (see Figure 1, middle panels). Each pixel in the resulting maps (a total of five maps for each land use class) reflected the percent of a given land use class within the kernel (i.e., if a 100 m square surrounding a given pixel is completely made up of agriculture, that pixel would be given a value of 100 for the agriculture variable). These distance calculations were then transformed using the eight transformations found in the R package “ResistanceGA” (named and illustrated in Figure 1, right panels; also see Peterman 2018). The genetic algorithm optimization method available in “ResistanceGA” would be computationally restrictive due to the large size of the study areas, so for purposes of tractability, we calculated transformed values using the “ResistanceGA” package with max=100 and shape=2.

Using the “gDistance” package (van Etten 2017) in R, we computed pairwise random-walk distance between individuals for each map, resulting in 40 distance calculations for each land use class. We also created a raster file that had a uniform pixel value of one to calculate a random-walk distance that would represent the geographic distance between points and could be used to test for isolation by distance (IBD). To remove the effect of geographic distance from our land use class random-walk calculations, we performed a series of simple linear regressions of the uniform pixel distance and each random-walk calculation using the “lme4” package in R (Bates et al. 2015). The residuals from these linear regressions were then used in model testing.

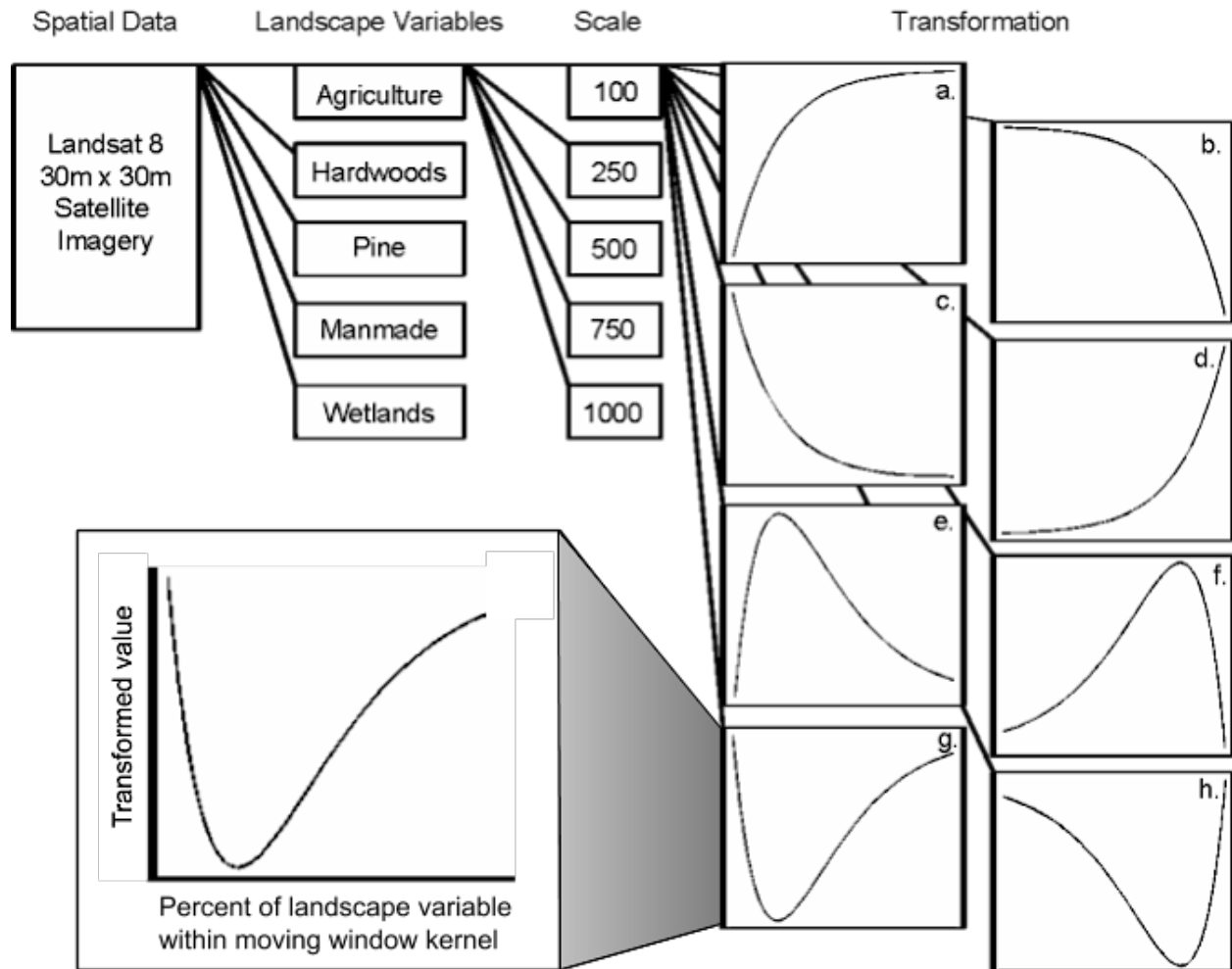


Figure 1. Depiction of landscape classification and optimization for scale and transformation.

Landsat 8 imagery was classified into 5 landscape variables. Each variable was tested for five scales (100, 250, 500, 750, and 1000 meters) and eight transformations (a. monomolecular, b. reverse monomolecular, c. inverse monomolecular, d. inverse-reverse monomolecular, e. ricker, f. reverse ricker, g. inverse ricker, and h. inverse-reverse ricker) and a linear relationship, for a total of 45 univariate tests per landscape variable. The transformation graphs show the relationship between the original resistance value (i.e., a value 0-100, indicating the percent of the given landscape variable within 100, 250, 500, 750, or 1000 meters) on the x-axis, and the new resistance value as a result of transformation on the y-axis, as is depicted in the inset.

Model Testing

To optimize each land use type for both transformation and scale, we ran a series of univariate maximum likelihood population effects models (MLPE). These linear random effects models account for the lack of independence between pairwise comparisons. This method was the most robust among seven regression-based model selection methods tested using inter-individual landscape genetic simulations (Shirk et al. 2018). Furthermore, in species distribution modeling simulations, generalized linear mixed models have been shown to be more transferable than those generated using machine learning and random forest methods (Wegner and Olden, 2012). Univariate models were ranked using corrected Akaike Information Criterion (AICc; Hurvich and Tsai 1989). The most strongly supported scale and transformation of each land use class (i.e., that with the lowest AICc score) was used for final model testing.

We tested several hypotheses of resistance with Maximum likelihood population effects (MLPE) models (Table 1). Each model included the geographic distance variable derived from a uniform raster, as well as a combination of land use variables. Models were then ranked using AICc (Table 1). Summaries of the best-fit models were examined to determine the sign of effect for each model component (i.e., each landscape variable). A positive sign of effect indicated that the variable resisted gene flow, whereas a negative sign of effect indicated the variable facilitated gene flow (Row et al. 2017).

Table 1. Maximum likelihood population effects models and AICc scores.

The lowest AICc scores for each forest are in bold.

Model Name	Variables Included	HSNF	BNF
Full Model	Geographic Distance, Agriculture, Manmade, Pine, Hardwood, Wetlands	29467	25809
Isolation by Distance	Geographic Distance	30089	26324
Modified Habitat	Geographic Distance, Agriculture, Manmade	29879	25990
Moderate Habitat	Geographic Distance, Pine, Agriculture	29467	25937
Forest Cover	Geographic Distance, Pine, Hardwood, Wetlands	29914	26192
Agriculture	Geographic Distance, Agriculture	29952	26018
Manmade	Geographic Distance, Manmade	29991	26249
Pine	Geographic Distance, Pine	30025	26269
Hardwood	Geographic Distance, Hardwood	29954	26209
Wetlands	Geographic Distance, Wetlands	29893	26104

Results

Genetic Analysis

Multilocus genotypes were produced from 113 of 114 individuals sampled in HSNF with 1.7% missing data, and 107 of 109 individuals sampled in BNF with 4.1% missing data. The remaining individuals (one individual in HSNF and two individuals in BNF) were excluded from the dataset because they failed to amplify at greater than two loci after repeated attempts. The locus *402* failed to amplify reliably in BNF, and was therefore removed from datasets in BNF but not HSNF. The locus *B8DRY* was found to be monomorphic in HSNF but not BNF, so it was removed from datasets in HSNF but not BNF. Loci within the HSNF dataset had 4-32 alleles, and loci within BNF had 10-29 alleles. Tests for departures from HWE using sampling locations with 9-11 individuals showed all loci were in HWE except one (*QWZ*) in HSNF and one (*43M*)

in BNF. Tests for linkage disequilibrium and null alleles found no linkage, and only one possibility of null alleles (*43M*) in BNF. *QWZ* was in HWE in BNF, and *43M* was in HWE in HSNF, and there was no indication of null alleles in *43M* in HSNF, therefore both loci were kept in the dataset. Mean allelic richness in HSNF was 14.5, and mean allelic richness in BNF was 18.9. STRUCTURE analysis supported $K = 1$ clusters (Evanno et al.'s 2005 method calculates a delta K value which by definition cannot be calculated for $K=1$, therefore we relied on the highest mean estimated log likelihood for each forest, which was at $K=1$). There was significant spatial autocorrelation within both forests (test for heterogeneity $p < 0.001$), with the x intercept at 7.38 km in HSNF and 16.28 km in BNF (Appendix Figures A1 and A2).

Landscape Analysis

Supervised classification of the Landsat 8 imagery and subsequent analysis with FRAGSTATS revealed HSNF and BNF have a number of similarities in the amount and distribution of landscape variables as well as a number of differences (Table 2). There are similar amounts of hardwood, manmade, and wetland areas in both forests, however there is approximately ten times more pine in HSNF than BNF. Patches of pine patches were more densely distributed across the landscape in HSNF (i.e., higher patch density) and they also had a higher correlation length, which is a measure of the distance an individual could travel and remain in a single patch when dropped in a random location and traveling in a random direction (Keitt et al. 1997; McGarigal et al. 2012). The amount of agriculture in HSNF was higher than in BNF, but the patches were at a similar density and the forests had similar correlation lengths. Areas containing manmade structures were considerably denser in BNF, however HSNF had a

higher average manmade correlation length, suggesting manmade structures in HSNF are mainly roads, and in BNF they are more commonly buildings and paved lots. Hardwood patches were denser in HSNF, but had a higher correlation length in BNF, and wetlands in BNF had a much higher correlation length than in HSNF.

Model Testing

The full model had the lowest AICc and was thus the best-fit model for both forests, indicating all of the tested landscape variables contribute to the genetic distances found in *P. mississippi*. The rank of model components differed between forests (Table 3), with agriculture and wetlands the top two components for both forests and the remaining variables contributing less to *P. mississippi* genetic distance variability. Agriculture, pine, and hardwoods had the same sign of effect in both forests, with agriculture and hardwoods resisting gene flow of *P. mississippi* and pine facilitating gene flow. In HSNF, manmade areas facilitated gene flow, whereas in BNF they resisted gene flow. The opposite was true for wetlands, which resisted gene flow in HSNF and facilitated gene flow in BNF (Table 4).

Univariate tests for scale and transformation resulted in few similarities between the forests (Figure 2). Only one landscape variable, pine, was optimized to the same transformation (inverse-reverse ricker). The presence of pine on the landscape resulted consistently in increased gene flow in both forest regions. In HSNF, gene flow was at its highest when ten percent of the area within 1000 m was comprised of pine. As the percentage of pine increased within the 1000

m area, the facilitation of gene flow decreased until it reached 80%, at which point it exponentially increased. The scale at which this occurred in BNF was 750 m.

Table 2. Comparison of landscape metrics for Holly Springs National Forest and Bankhead National Forest.

Habitat amount (km²) Patch Density (number of patches per 100 ha), Correlation Length, Clumpiness Index, Patch Cohesion, Aggregation Index for Holly Springs National Forest (grey rows) and Bankhead National Forest (white rows).

Landscape Type	Amount of Habitat	Patch Density	Correlation Length	Clumpiness Index	Patch Cohesion	Aggregation Index
Hardwood	1339.41	14.04	353.68	0.59	93.24	68.73
	1382.44	9.96	912.92	0.59	97.52	79.60
Pine	2230.41	17.22	562.51	0.61	95.53	76.65
	236.40	6.41	211.13	0.69	89.43	71.51
Agriculture	958.83	8.03	295.24	0.69	92.13	74.53
	271.55	4.10	258.56	0.73	92.09	75.42
Manmade	291.75	1.42	5511.04	0.53	98.80	55.50
	311.57	10.66	3838.83	0.51	98.61	56.48
Wetlands	245.30	11.59	91.31	0.41	73.57	43.67
	441.08	17.71	279.73	0.55	89.24	62.20

Table 3. Rank and model coefficients of landscape variables in most supported Maximum likelihood population effects model for each forest region.

HSNF		BNF	
Wetlands	0.91	Agriculture	1.14
Agriculture	0.69	Wetlands	-0.48
Hardwoods	0.65	Manmade	0.47
Manmade	-0.43	Pine	-0.40
Pine	-0.25	Hardwoods	0.26

Table 4. Comparison of scale, transformation, and sign of effect for Holly Springs National Forest and Bankhead National Forest.

Results for Holly Springs National Forest are in grey and results for Bankhead National Forest are in white. A negative sign of effect indicates the variable facilitates gene flow, and a positive sign of effect indicates the variable resists gene flow. For example, in Holly Springs National Forest, gene flow among *P. mississippi* populations is restricted by the presence of hardwoods up to 500 m away. This resistance is at its lowest when hardwoods comprise 20% of the 500 m kernel, and at its greatest when hardwoods comprise 100% of the 500 m kernel.

Landscape Type	Scale (m)	Transformation	Sign of effect
Hardwood	500	Inverse Ricker	+
	100	Inverse-Reverse Ricker	+
Pine	1000	Inverse-Reverse Ricker	-
	750	Inverse-Reverse Ricker	-
Agriculture	500	Inverse-Reverse Ricker	+
	500	Inverse Ricker	+
Manmade	250	Inverse Ricker	-
	1000	Ricker	+
Wetlands	1000	Inverse Ricker	+
	100	Monomolecular	-

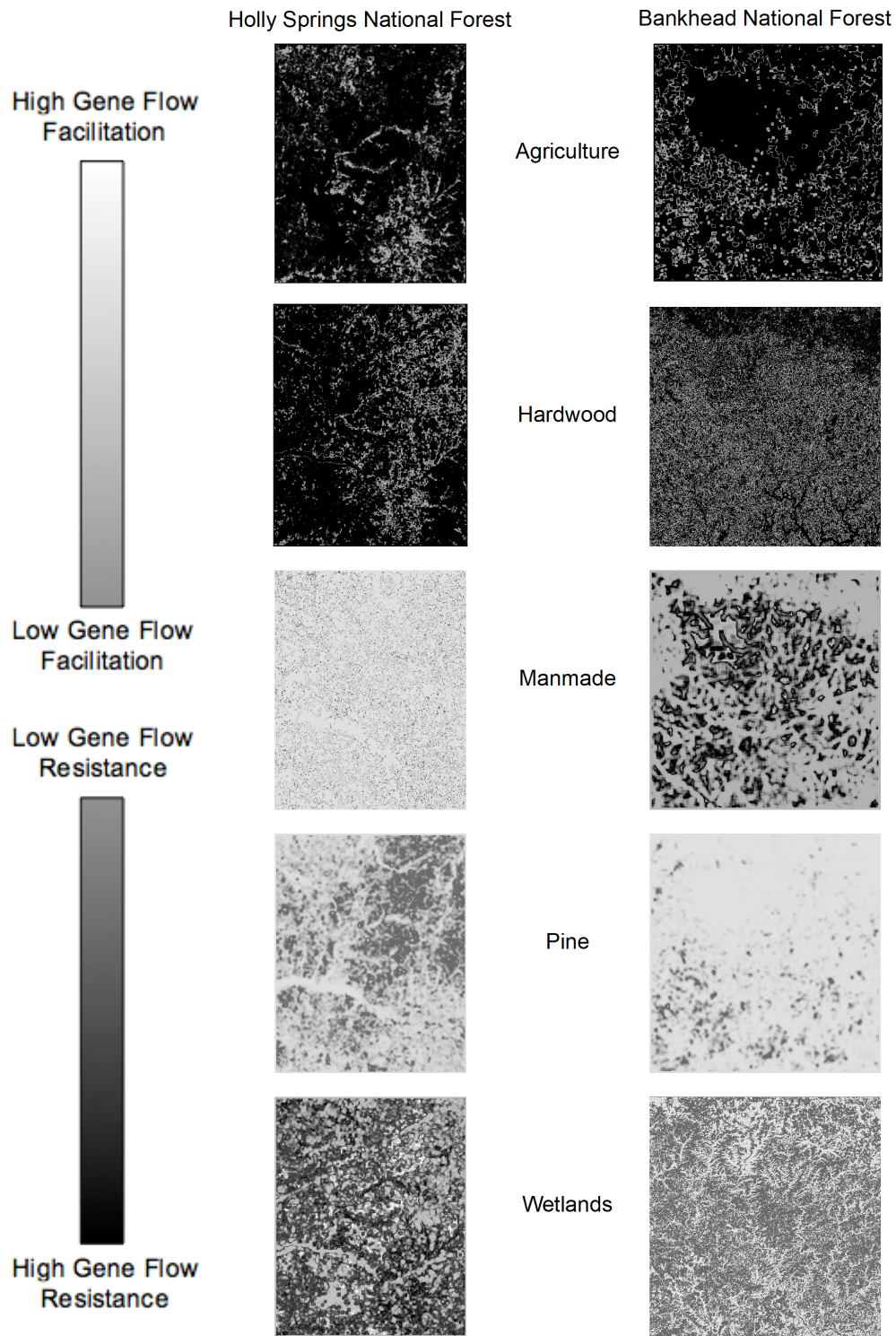


Figure 2. Comparison of the facilitation or resistance to gene flow created by land use types in Holly Springs National Forest versus Bankhead National Forest.

Maps illustrate the optimized scale, transformation, and sign for each land use type to give a visual representation of the effect of each land use type on gene flow of *P. mississippi*.

Only one landscape variable, agriculture, was optimized at the same scale (500 m). Agriculture resulted in a resistance to *P. mississippi* gene flow in both study regions. The transformation, or function of the relationship between gene flow and the amount of agricultural land, differed between forest regions. In HSNF, small percentages of agriculture created the highest resistance (i.e., the area within 500 m comprised of 20% agriculture) and the lowest amount of resistance was seen when 80% of the surrounding area was agriculture. Conversely, in BNF, the lowest resistance to gene flow occurred when 20% of the surrounding landscape was agricultural.

The largest difference in scale was found in wetlands, with wetlands in HSNF impacting gene flow at a scale of 1000 m, but wetlands in BNF impacting gene flow at a scale of 100 m. Wetlands in HSNF correlated with resistance to gene flow, with increasing resistance from 20–100% wetlands within 1000 m. In contrast, wetlands in BNF correlated with facilitation of gene flow at a scale of 100 m with increasing facilitation with an increasing percentage of wetlands within a 100 m kernel.

The manmade landscape variable showed the most drastic difference in transformation, with HSNF optimized to an inverse ricker transformation and BNF optimized to a ricker transformation. In HSNF, manmade structures facilitate gene flow, whereas in BNF, manmade structures correlate with resistance to gene flow. However, facilitation is at its lowest when manmade structures comprise 20% of the landscape in HSNF, and resistance is at its highest when manmade structures comprise 20% of the landscape in BNF.

Hardwoods created resistance to *P. mississippi* gene flow in both forest regions, however both the scale and function (i.e., transformation) of their effect was different. The scale of effect in HSNF was at 500 m, with increasing resistance from 20-100% of the immediate area consisting of hardwoods. In BNF, the scale of effect was at 100 m, showing decreasing resistance from 0–80% hardwoods.

Discussion

In this study, landscape genetics models for the Mississippi slimy salamander were evaluated in each of two forest regions that, despite being geographically separated by ~190 km, are nonetheless qualitatively similar in many respects (e.g., forest types, land uses, management, climate, and precipitation). This metareplication allowed us to understand the extent to which landscape genetics inferences are transferable to neighboring regions. Indeed, due to the general similarity between the two forest regions, the present study represents a case where transferability is potentially quite high. In both forest regions, the most strongly supported MLPE model included all five of the landscape variables under consideration, indicating they all influence *P. mississippi* gene flow in non-negligible ways, even if the magnitude of influence is weak for some. We found that the rank ordering of variable effects was different between forests, and so generally speaking, the notion of a single landscape genetic model that is broadly applicable across the species' range was not supported. Notably, wetlands and agriculture were consistently at the top of the rankings for both regions. Furthermore, some variables affected gene flow in the same way across the two forest regions (e.g., agriculture generated resistance to

gene flow), suggesting some generality the responses of *P. mississippi* individuals to characteristics of the intervening matrix. That said, there was also a number of variables, including some of large effect (e.g., wetlands), which differentially affected gene flow, indicating context-dependent responses that may make extrapolation of landscape genetics models to as-yet unstudied areas difficult. Below, we discuss the variables in descending order of rank of effect, and compare our findings with those of other relevant studies. We close by considering some limitations of our work, and point to future directions for understanding the effects of landscape genetics model optimization (e.g., choice of geographic scale, transformation of resistance distances) upon generating broadly applicable inferences about how individuals perceive and move through their habitat.

Based on our data, agriculture was ranked as one of the top two most influential variables affecting gene flow in the best-fit landscape genetics models for *P. mississippi* in both forest regions. We found that agricultural areas caused resistance to gene flow at a relatively intermediate spatial scale of 500 m. Notably, the most prominent agricultural practices within both study regions include a rotation of corn, wheat, and soybeans, as well as cotton and sweet potatoes. In all of these cases, the crops are grown as monocultures, and harvesting/planting times are such that there is high probability of bare earth during spring and fall, which are the most active seasons for *P. mississippi* in terms of dispersal of individuals (Salmerón et al. 2016; Petranka 1998; S. Burgess personal observation). The lack of vegetative cover and increased ground disturbance associated with planting and harvest activities may explain the resistance to gene flow caused by agricultural areas. Abundance surveys of plethodontid salamanders have shown a direct relationship between individual salamanders and the amount of herbaceous cover

(Riedel et al. 2008) and recent landscape genetics analyses of marbled newts have shown an increase in population structure due to agricultural areas (Costanzi et al. 2018).

Like agriculture, wetlands ranked in the top two most influential variables. However, the effects of wetland configuration on gene flow differed between the two forest regions; gene flow was resisted at a large scale (1000 m) in HSNF, whereas gene flow was facilitated at a much smaller scale (100 m) in BNF. These contrasting outcomes may be due to differences in the shape and connectedness of wetland patches across each forest region. For instance, in BNF, the correlation lengths of wetland patches (i.e., a metric of the amount of time an individual can move forward from a random starting point in a random direction and stay within the same patch) are considerably higher than in HSNF. In this context, simulations by Cushman et al. (2011; 2013) are particularly relevant, as these authors explored how the distribution of landscape variables within a study area impacts their effect on gene flow, finding the most prominent effects from differences in correlation length and patch cohesion, (i.e., a metric indicating the physical connectedness of patches). Thus, our data suggest that in BNF, individual *P. mississippi* can move relatively large distances without exiting wetland areas, but not in HSNF. While the notion of long distance dispersal in salamanders is counter to the traditional view of extreme philopatry in these taxa, it is worth noting that for most species, dispersal distributions are leptokurtic. This has been demonstrated by occasional dispersal greater than 400m by the spring salamander, *Gyrinophilus porphyriticus* (Lowe 2010), and the Near Eastern fire salamander, *Salamandra infraimmaculata* (Bar-David et al. 2007). Overall, our data underscore the notion that if wetlands are well connected, they facilitate gene flow in terrestrial

salamanders such as *P. mississippi*, however isolated wetlands, such as those found in HSNF, may lead to restriction of gene flow.

Our results showed that the magnitude of influence on gene flow of hardwood, manmade, and pine landscape variables were consistently ranked relatively low in the best-fit MLPE models for both forest regions. While hardwoods consistently generated resistance to gene flow in both forest regions, the function of resistance (i.e., the transformation) differed. Specifically, in HSNF, there was a positive relationship between amount of hardwoods and amount of resistance, whereas in BNF, the reverse was true. As with wetlands, the correlation lengths of hardwood patches in BNF were higher than in HSNF. One possible explanation for these relationships is that *P. mississippi* typically reside in cool, moist, bottomland hardwood forests (Petranka 1998) and thus may be unprompted to leave. Smith and Rissler (2010) reported that hardwood dominated understories were characteristic of "pristine" habitat for terrestrial herpetofauna in Talladega National Forest, Alabama, and the population genetic effects of local philopatry have been reported for plethodontid salamanders, in continuously forested habitat (e.g., *P. cinereus*; Cabe et al. 2007). Thus, in the case of *P. mississippi*, the reduction in gene flow, and associated apparent "resistance" generated by hardwoods in the best fit landscape genetic models, should not be equated to that caused by agriculture. Indeed, Richardson et al. (2016) cautioned that even when a set of landscape variables each show evidence of resistance, the underlying reasons may differ ecologically.

In HSNF, we found that manmade structures facilitated gene flow among *P. mississippi*, whereas in BNF they correlated with resistance to gene flow. Notably, patches of manmade

structures in HSNF are considerably less dense, but have a greater correlation length, suggesting their distribution across the landscape is more linear (i.e., more roads than buildings and paved lots). Studies of salamander gene flow have found variable responses to roads, showing both resistance to gene flow (Marsh et al. 2008) and no effect (Purrenhage et al. 2009). Our analysis was unique due to the inclusion of manmade structures, however our results support separating roads and other manmade structures in the future to differentially determine the effect of each land use type.

Although pine ranked as one of the least influential landscape variables in each best-fit MLPE model, it did have the most consistent effect across the two forest regions (i.e., same sign, scale, and transformation). This consistent facilitation of gene flow by a landscape type that is less than ideal habitat for *P. mississippi* (Petranka 1998) may seem counter to predictions based on the species' natural history. However, increased dispersal—and by extension, gene flow—through moderately hostile habitat has been reported for ambystomatid (Wang et al. 2009) and plethodontid (Peterman et al. 2014; Prunier et al. 2014) salamanders. In each of the two forest regions studied here, pine is typically found immediately adjacent to hardwoods and, as a consequence of being evergreen, provides year-round canopy cover. Furthermore, a considerable amount of the pine habitat found in both forest regions is routinely burnt via prescribed low-intensity fire (U.S. Department of Agriculture Forest Service 2004, 2012), resulting in a reduction of pine litter, which has been experimentally shown to increase the movement of an ambystomatid salamander (*Ambystoma talpoideum*, Moseley et al. 2004). Thus the combination of these factors—close proximity to ideal habitat, protection by year-round canopy cover, and

potential to increase movement—may explain the consistent increase in gene flow found in response to pine habitat.

Due to the nature of metareplications, there are a number of potential limitations that arise both during analysis and in the interpretation of results. Because the goal of metareplication is to compare the effect of each landscape variable in multiple areas, researchers must include environmental and landscape variables that have both the potential to affect the genetic structure of their study organism (Keller et al. 2014) and are also present in all study locations (Short Bull et al. 2011; Castillo et al. 2016; Vergara et al. 2017). When results are intended to apply to conservation throughout a species range (Row et al. 2015) researchers may focus on the inclusion of a smaller number of landscape variables in an attempt to strike a balance between the number of parameters evaluated and the transferability of model inferences. By focusing on a smaller number of landscape variables, researchers may fail to identify a landscape feature that affects gene flow. There is also a potential source of error in the interpretation of metareplication results. Because metareplications must, by definition, occur in separate geographic areas, they inherently include the potential for erroneous conclusions due to unidentified phylogeographic breaks between study areas. Divergent lineages can be cryptic, abrupt, and even counter to morphologic differences (e.g., in two *Desmognathus* species identified by Jones and Weisrock 2018). If a phylogeographic break exists, any inferred location-specific differences between study regions (e.g., response to a particular landscape variable) may instead be due to deeply divergent lineages and separate evolutionary histories. In the present study, we have prioritized the optimization of a limited number of landscape variables to maintain the ability to compare their effects across forest regions while attaining detailed information about the scale and

function (i.e., transformation) of those effects. While the possibility of an unidentified phylogeographic break exists between *P. mississippi* in our two forest regions, both HSNF and BNF lie well within the range delineated for *P. mississippi* by Highton (1989).

Management Implications

Through metareplication, we have obtained information about the relationships between *P. mississippi* and their environment that would not have been apparent when using a single study area. This information can be used to generate management recommendations for the species in as-yet unstudied locations. The consistent resistance to gene flow by agriculture, coupled with the consistent facilitation of gene flow by pine, indicates that land managers overseeing multi-use areas could increase *P. mississippi* gene flow by prioritizing silviculture over agricultural leasing. Our results also indicate that the connectivity of wetland patches (as seen in our analyses as a high correlation length) is an important factor in their ability to facilitate *P. mississippi* gene flow. Thus, managers should focus efforts on improving the connectivity of wetlands, potentially through targeted restoration.

CHAPTER 3

THE EFFECT OF GEOGRAPHIC SAMPLING SCALE ON LANDSCAPE GENETICS INFERENCE FOR THE SLIMY SALAMANDER (*PLETHODON MISSISSIPPI*)

Abstract

Within landscape genetics, the effects of sampling density and study area size upon inferences are largely unknown. A common recommendation is that sampling locations be placed no further apart than average individual dispersal distance, leading to a small, unrepresentative study area or a logistically challenging number of sampling locations. We tested the effects of sampling density and study area size on landscape genetics models for *Plethodon mississippi* in Mississippi, USA, via comparative analysis of nested datasets that differed in sampling density and study area size. Genetic distances among individuals were divided into datasets representing dense sampling across a small study area, sparse sampling across a small study area, and sparse sampling across a large study area. These datasets were used in models that assessed the influence of land use classes on resistance or facilitation of gene flow. Wetlands were a significant contributor to genetic distance, correlating with gene flow resistance in all datasets. Correlations between gene flow facilitation and manmade structures, and gene flow resistance and hardwoods were also consistent across datasets. Small study areas resulted in correlation between pine and gene flow resistance, whereas a large study area found correlation

to gene flow facilitation. The relationship between agriculture and gene flow appeared to be affected by both study area and sampling density. Differences found when using study areas of different sizes and patch configurations suggest these factors may affect model inferences. Differences found when using dense versus sparse sampling suggest that short distance dispersers may interact with their environment differently than long distance dispersers. Metareplication in areas with different patch configuration combined with reanalysis of "thinned" subsets of original data, mimicking different sampling densities, may capture these important differences.

Introduction

In the field of landscape genetics, features of the environment that may affect dispersal and gene flow are represented by a suite of ecological distances (i.e., predictor variables) that are compared to corresponding genetic distances (i.e., the response variable) between individuals or local populations (Manel et al. 2003). The results of this comparison are intended to identify abiotic or biotic characters of the landscape that influence genetic connectivity, such as riverine or road barriers (Hartmann et al. 2013), agricultural land use practices (Goldberg and Waits 2010; Prunier et al. 2014; Costanzi et al. 2018), and the spatial configuration of preferred or non-preferred habitat (Vergara et al. 2017). Landscape genetics analyses have been conducted on many species and can be applied broadly across different landscape settings (Storfer et al. 2010). Accordingly, these studies have spanned a broad range of spatial scales [e.g., a study area <40 km² for the Natterjack toad, *Epidalea calamita* (Cox et al. 2017), up to the entire state of

Wyoming, USA (250,000 km²), for the greater sage-grouse, *Centrocercus urophasianus* (Row et al. 2015)]. Given that spacing between locations where DNA samples are collected can affect inferences about which ecological variables have the greatest impact on gene flow (Richardson et al. 2016), sampling density and size of the study area are important considerations.

Decisions about geographic sampling scale are made *a priori*, based in part on an understanding of species' dispersal ability, whereby species with short dispersal distances are often sampled more densely than species for which long distance dispersal is common. Given that fieldwork can be labor-intensive and expensive, in the case of species with short distance dispersal, researchers may be faced with the decision to either increase the spacing between sampling locations, or conduct their study within a smaller area. However, smaller study areas may fail to incorporate the level of environmental heterogeneity needed to answer research questions. For instance, to test the effect of roads on gene flow, a study area must encompass a moderate to large number of roads, preferably of varying sizes (Keller et al. 2014; Richardson et al. 2016)—a requirement that may not be satisfied by a small study area. Accordingly, there is a need for investigations that explicitly evaluate the impacts of alternative spatial arrangements of sampling sites, using the same focal species and landscape setting, such that the direct effects of geographic sampling scale on outcomes from landscape genetics can be examined.

Researchers have cautioned that landscape genetics inferences based on sparse genetic sampling (i.e., where average distances among locations far exceed typical movement of individuals) may fail to capture relationships between abiotic features of the landscape and gene flow that are unique to short distance dispersal events (Angelone et al. 2011). Angelone et al.

(2011) analyzed the relationship of 16 landscape variables and gene flow among populations of European tree frog (*Hyla arborea*). To assess the impact of different geographic sampling scales, the authors subdivided pairwise comparisons of breeding ponds into geographic distance classes, and analyzed these datasets separately. This tiered analysis resulted in different ecological predictor variables being identified as causing resistance to gene flow for each distance class examined. The outcome was considered to be consistent with a scenario where individuals exhibiting short distance dispersal were affected by the presence of rivers or lakes, whereas geographic distance, wetlands, hedgerows and the density of forests more strongly affected long distance dispersers (Angelone et al. 2011). Sparse sampling can also result in a weaker relationship between genetic distance and the ecological variables that impact gene flow due to the greater potential for stochastic events (e.g., local extreme weather, invasive or predatory species interactions, disease spread) to occur between sampling sites (Epperson et al. 2010). For example, in a wetland grasshopper (*Stethophyma grossum*), Keller et al. (2013) found the greatest model fit (measured by strength of correlation between genetic and ecologically-informed geographic distances) was obtained when only including pairs of sampling locations that were within close proximity (i.e., up to 3 km apart; the threshold for minimum population connectivity in that study system). Those authors suggested that the decrease in model fit when examining widely separated populations may occur because the rarity of long distance dispersal events reduces the ability to detect a relationship between ecological variables and long distance movements. Both groups of researchers recommended that sampling locations be spaced no further than the average individual dispersal distance of the focal species.

Just as there may be negative consequences for landscape genetics inferences when sparse sampling is used, this may also be true for extremely fine scale sampling. The typical logistical tradeoff in which dense sampling is coupled with a smaller overall study area may create a situation where a representative range of values of one or more ecological predictor variables (e.g., the magnitude of differences between high versus low amounts of agriculture) are not captured by the study design, such that it becomes difficult to identify an environmental variable's true, overarching impact on gene flow (Keller et al. 2014). Haran et al. (2017) repeatedly subsampled their dataset of individual-based genetic distances among pine sawyer beetles (*Monochamus galloprovincialis*) from the Iberian Peninsula, and found significant relationships between gene flow and environmental variables were more likely to be detected when using larger study areas. When testing the relationship between environmental variables and genetic distance in over 30,000 alternative demarcations of a study area (220–1,000 km in diameter), the number of significant relationships between gene flow and high elevation, cooler temperatures, and pine forests was highest when study areas were large (1,000 km diameter), and the relationships between gene flow and cooler temperatures was highest when study areas were smaller (600 km diameter.) While this study highlights the need to include large study areas, its findings are also consistent with Angelone et al. (2011) and Keller et al. (2013) in supporting the idea that the impact of environmental variables on gene flow may be scale-dependent. Thus, there are reasons for concern regarding overly small study areas, as these studies may fail to detect significant relationships between ecological predictor variables and gene flow that do exist.

In the present study, we explored the impact of sampling density and study area size on landscape genetics inferences for a low-mobility amphibian—the Mississippi slimy salamander, *Plethodon mississippi* (Highton 1989). We approached this by first sampling at two contrasting densities, and then by reanalyzing a subset of the empirical data. This salamander is found within the bottomland hardwood forests and wet pine-woods of Mississippi, Alabama, and Tennessee, USA (Petranka 1998). The species develops terrestrially without the need to disperse to aquatic environments for reproduction. In northern Mississippi, *P. mississippi* is distributed continuously throughout Holly Springs National Forest (HSNF), a 630 km² federally managed forested area that contains a mosaic of hardwood forests, manmade structures, agricultural fields, roads, and pine plantations (U.S. Department of Agriculture Forest Service 2012). In order to determine the effect of this mosaic on gene flow among populations of *P. mississippi*, the sampled area must encompass a representative portion of the forest. However, given that individuals typically disperse less than 92 meters in their lifetime (Wells and Wells 1976), spanning such a large area with locations spaced no further than average dispersal distances is intractable. Thus, this is a suitable model for investigating the relationship between sampling density and study area size. Here, we assessed the relationship between five ecological predictor variables and *P. mississippi* gene flow using genotypic data from eight microsatellite loci. The effects of sampling density on landscape genetic inferences were examined using three different sampling schemes: sparse sampling across a large area, sparse sampling across a small area, and dense sampling scale across the same small area (Figure 3). Based on this, we addressed the following two questions: 1) Does sparse sampling fail to identify relationships between ecological predictor variables and gene flow that are identified using dense sampling? 2) If differences exist between inferences

obtained from the two contrasting sampling densities, are they a consequence of sampling density (i.e., sparse versus dense) alone, or does study area size also play a role?

Methods

Study Design

To examine the effect of sampling density on landscape genetic inferences, hierarchically nested datasets were created. These datasets contained combinations of two sampling strategies and two study area sizes. The first, herein referred to as the dataset generated using “sparse sampling across a large study area,” consisted of 19 sampling locations placed approximately seven kilometers apart distributed evenly across HSNF—a forest region that spans approximately 45 km x 70 km (Figure 3A). The second, herein referred to as the dataset generated using “dense sampling across a small study area”, consisted of 14 sampling locations placed approximately three kilometers apart across a 16 km x 16 km area nested within the larger HSNF study area that was similar in land use composition (Figure 3B). The third dataset, herein referred to as “sparse sampling in a small study area,” included only the “dense sampling in a small study area” pairwise comparisons between individuals greater than seven kilometers apart (Figure 3C).

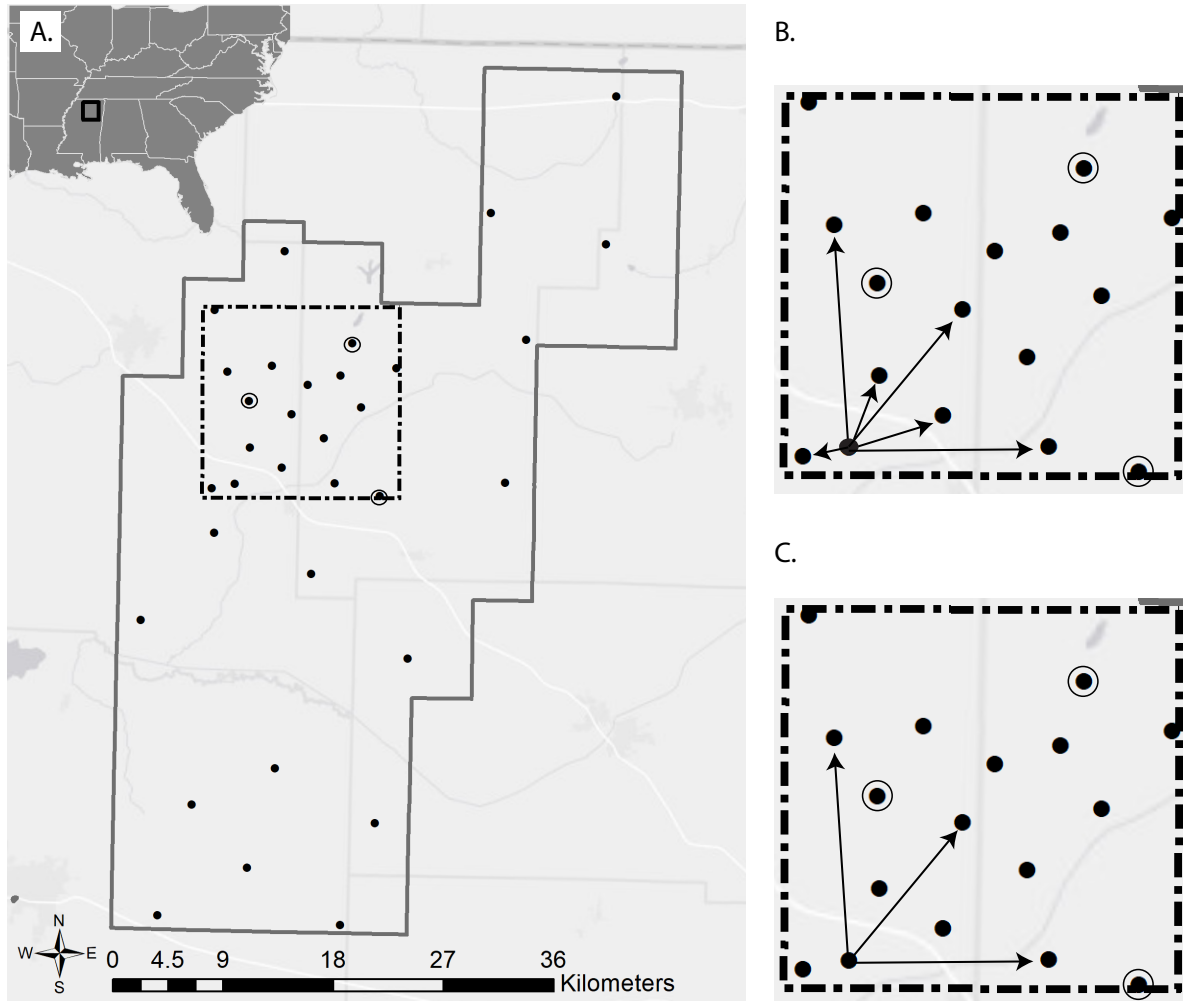


Figure 3. Sampling locations of *P. mississippi* within Holly Springs National Forest (HSNF), Mississippi USA.

A) In the sparse sampling across a large study area (630 km²), 19 sampling locations (black dots) were spaced approximately 7 km apart across the entirety of HSNF. Within the small study area (256 km²) demarcated by a dashed box, 14 sampling locations were spaced approximately 3 km apart. The circled sampling locations within the small study area were also part of the large study area dataset (Inset: map of southeastern USA showing location of HSNF). B) The dense sampling across a small study area included pairwise genetic distances between individuals from all sampling locations. C) The sparse sampling across a small study area only included pairwise genetic distances from individuals that were > 7 km apart.

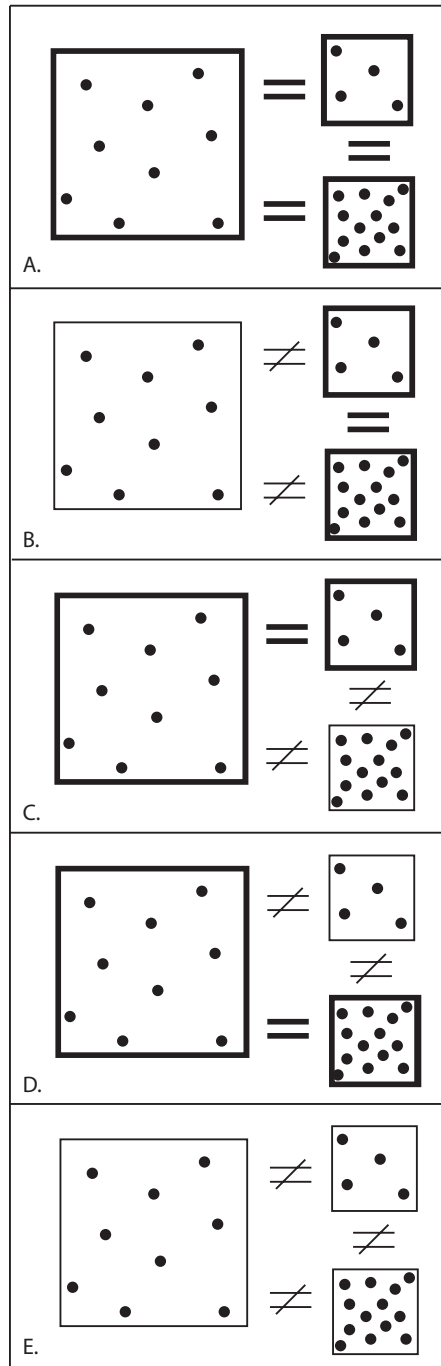


Figure 4. Different combinations of sampling density and study area size, and hypothetical outcomes relating to similarity of landscape genetics inferences among datasets.

A) similar outcomes are obtained for all three datasets. B) sparse and dense sampling across a small study area yield similar outcomes, but differ from the sparse sampling across a large study area. C) sparse samplings across a large and small study area yield similar outcomes, but differ from dense sampling across a small study area. D) sparse sampling across a large study area and dense sampling across a small study area yield similar outcomes, but differ from sparse sampling across a small study area. E) all models differ. A combination of the effects seen in B) and C) may be the cause of D) or E).

Comparison of the three datasets can be used to assess support, or lack thereof, for the research questions outlined above. If the effects seen for a given ecological predictor variable are consistent across these three datasets (e.g., Figure 4A), this would indicate that sparse sampling may be adequate to detect this relationship between the ecological predictor variable and gene flow, and dense sampling may not be necessary. If the sparse datasets (both in large and small study areas) result in a consistent effect, but that effect is different from that found using dense sampling in a small study area (Figure 4B), it would suggest that sparse sampling may fail to detect relationships between ecological predictor variables and gene flow among *P. mississippi* populations. If the datasets in small study areas (both sparse and dense sampling) result in similar effects that are different from those generated using sparse sampling in a large study area (Figure 4C), this would imply that datasets using a large study area may detect relationships between ecological predictor variables and gene flow among populations that go undetected in small study areas. Additionally, there may be similarities between the datasets using sparse sampling in a large study area and dense sampling in a small study area, or all of the results may differ (Figure 4D and E, respectively). These similarities and differences may be due to a combination of the effects described above.

Geographic Sampling Schemes

To determine where to place the boundaries of small study area so that it best represented the land use composition of the larger, forest-wide study area, we first created land use maps using NASA Landsat 8 satellite imagery. Spectral bands 3, 4, 5, and 6 of the 30 x 30 m images were classified into six land use classes (agriculture, hardwood forest, pine forest, manmade

structures, water bodies, and wetlands) using a supervised classification method in ERDAS Imagine 2014 (Hexagon Geospatial, Norcross, GA, USA). Flowline and wetland shapefiles obtained from the US Fish and Wildlife Service National Wetlands Inventory and roads shapefiles derived from the US Forest Service Motor Vehicle Use Map (developed using data from the US Census Bureau) were overlaid onto the classified image using the raster calculator tool in ArcGIS 10.2.2 (ESRI 2011) to classify roads, wetlands, and waterways that may be difficult to identify using aerial classification. To determine the most representative location for the small study area, a 16 km x 16 km square polygon shapefile was created in ArcGIS, and was then moved across the classified raster file until the land use class percentages were within 6.3% of the percentages found in the larger study area (Appendix Table A4). Three of the large study area sampling locations also fell within the small study area, thus individuals from these locations were included in both large study area and small study area datasets.

Genetic Sampling

At each sampling location, tail tissue was sampled from at least five *P. mississippi* individuals following procedures approved by University of Mississippi IACUC approval #15-020 and Mississippi Department of Fish and Wildlife Permit #0324164, and then stored in 95% ethanol. In total, tissues were collected from 184 individuals from 33 sampling locations within HSNF.

Genetic Analysis

Genomic DNA was extracted using a Qiagen DNeasy Blood and Tissue Kit (Valencia CA, USA), following the manufacturer's recommendations. Eight microsatellite loci reported by Spatola et al. (2013) were used to genotype individuals (see Appendix and Tables A1, A2, and A3 for amplification conditions, and allele-calling approaches). One individual that failed to amplify at more than one locus after multiple attempts was removed, leaving 183 individuals with multilocus genotypes. Given that *P. mississippi* is a continuously distributed species for which discrete local populations are not apparent within HSNF, we grouped individuals in three ways for the purpose of basic validation of molecular marker inheritance patterns. First, the entire dataset was grouped as a single, panmictic population. Second, three sampling locations (two within the large study area dataset, and one in both large and small study area datasets) where tissue from 9–10 individuals was collected, were each used to represent local populations. Third, the entire dataset was divided into putative populations by grouping sampling locations 20 km apart (past the extent of spatial autocorrelation as determined by our semivariogram; see Results) into eight different populations. Each of these three grouping schemes for designating putative populations were used to test for null alleles and departures from Hardy-Weinberg Equilibrium, using the package “PopGenReport” (Adamack and Gruber 2014) in R (R Core Team 2019). MICRO-CHECKER v. 2.2.3 (van Oosterhout et al. 2004) was used to test for departures from linkage equilibrium. “PopGenReport” was also used to calculate percent missing data, number of alleles per locus, and mean allelic richness across the total dataset.

We conducted analyses using an individual-based genetic distance measure (Shirk and Cushman 2014). We conducted a principle components analysis (PCA) with 64 axes of ordination using the R package “ade4” (Dray and Dufour 2007), given that this approach has been shown to out-perform other individual-based genetic distance measures when sample sizes and genetic structure are low (Shirk et al. 2017). We then calculated pairwise Euclidian distance between individuals using the 64 PCA axes. Pairwise genetic distance was calculated between all sampled individuals (from both large and small study area datasets) to assess the overall variation of genetic distance over geographic space. To determine whether the relationship between genetic distance and straight line geographic (Euclidean) distance was stronger for different distance classes, we regressed geographic distance against genetic distance using the lme4 package (Bates et al. 2015) in R for three nested groups: all pairs of individuals less than 10 km apart, all pairs of individuals less than 20 km apart, and all pairs of individuals. Nested group sizes were chosen so that the smallest group, 10 km, would include comparisons from both sparse and dense sampling efforts. Additionally, to assess evidence for spatial autocorrelation, determine if genetic differentiation occurs at multiple spatial scales (Wagner et al. 2005), and allow for visual representation of the geographic extent of any spatial autocorrelation, a semivariogram was created from genetic and geographic Euclidean distances using 52 distance classes with a distance interval of 1.5 km in the “phylin” package (Tarroso et al. 2019) in R. Distance class size was designated as smaller than the shortest distance between observations (i.e., 3 km) while minimizing the number of bins that lacked observations (50 of 52 bins contained observations). If the relationship between geographic and genetic distance differed across distance classes, we would expect this to be evident from different slopes for the nested regressions, and multiple plateaus within the semivariogram. For the remaining analyses,

pairwise genetic distance was calculated separately for the large and small study area sampling locations.

Landscape Analysis

The classified land use raster (see *Geographic Sampling Schemes*, above) was used to create a series of maps for each land use class using a square moving window analysis in FRAGSTATS v. 4.2 (McGarigal et al. 2012). The size of square moving windows were designated using the length of a side, thus a 250 m moving window represents a 0.0625 km² area. The value of each pixel within the map was calculated using the function PLAND, which determines the percent of the window that contains a given land use class (i.e., in the case of a 250 m window if all but a small 80 m x 80 m area consisted of pine, the pixel value would be 90). Five maps were created for each land use class, with moving windows of 100, 250, 500, 750, and 1000 m. To test for non-linear relationships between percent land use and gene flow, each moving window map was then transformed into eight different maps (named and illustrated in Figure 5) using the R package “ResistanceGA” (Peterman 2018). For purposes of computational tractability, we calculated all transformations using a max=100 (maximum resistance of 100) and shape=2 (an indicator of the shape of each function, unique to “ResistanceGA”).

Pairwise random-walk distances between individuals within each dataset (i.e., large study area or small study area) were calculated using each of the transformation maps (i.e., 45 maps for each land use class) using the R package “gDistance” (van Etten 2017). A raster file with a

uniform pixel value of one was also created to calculate random-walk distance on a homogenous landscape. The latter was used to test the effects of straight-line geographic distance on genetic distance. A series of linear regressions were performed between random-walk distances using the R package lme4 to assess correlation between land use classes. To isolate the effect of each land use class on gene flow, the random-walk distance for each map was regressed against the homogenous landscape distance using a simple linear regression, thereby removing the effect of geographic distance. The residuals from these simple linear regressions were used in model testing.

Model Testing

Three different datasets were analyzed (see *Study Design*, above). For each of these, we optimized for the best-fit scale (i.e., 100, 250, 500, 750, or 1000 m) and for the best-fit transformation (see Figure 5) of each land use class using a series of maximum likelihood population effects (MLPE) models. These models are a form of random effects model that account for the lack of independence introduced by including multiple individuals per sampling location, and for individual-based comparisons, MLPE models are considered robust (Shirk et al. 2018). Pairwise genetic distances between individuals that were collected from the same sampling location were removed to prevent skewing models. For each land use class, models were ranked using the corrected Akaike Information Criterion (AICc; Hurvich and Tsai 1989). Models with the lowest AICc score were considered to have the best fit, and were used in all subsequent analyses.

Using the optimized scale and transformation for each land use class, several multivariate models were generated to test combinations of land uses that contribute to variance in genetic distance (Table 5). These MLPE models tested pairwise genetic distance against straight-line geographic distance (generated using the random-walk across a homogenous surface) and a combination of optimized land use class distances. Models were ranked using AICc, and the best-fit models were examined to determine the sign of effect for each land use class, where land use classes with negative signs of effect facilitate gene flow, and those with positive signs of effect restrict gene flow (Row et al. 2017).

Results

Genetic Analysis

Total missing genotypic data for 183 individuals was 1.8%. A comparison of the results from our three alternative population grouping strategies showed that although loci *QWZ* and *241* showed evidence of homozygote excess in some putative populations, these loci did not consistently depart from HWE across all grouping strategies. Similarly, while the same two loci showed the possibility of null alleles, the frequency of potential null alleles was very low (i.e., <0.1 , as calculated in PopGenReport using the method outlined by Brookfield 1996). Thus, we retained all loci. Based on all 183 individuals, the number of alleles per locus ranged from 6–32, with a mean of 16.5 alleles per locus. The relationship between geographic and genetic distance

was significant for only the regression containing individuals less than 10 km apart (individuals less than 10 km apart: slope=0.07, adjusted $R^2=0.002$ $p=0.004$, individuals less than 20 km apart: slope=0.01, adjusted $R^2=-0.00005$, $p=0.78$, all pairs of individuals: slope=0.0006, adjusted $R^2=0.0001$, $p=0.17$). Consistent with these results, the semivariogram plateaued only once at 10 km (Figure 8), indicating only one scale over which spatial autocorrelation occurred.

Landscape Analysis and Model Testing

Tests for non-independence of predictor variables revealed a strong positive correlation between the wetland and water land use classes. Because the primary habitat for *P. mississippi* is wetland areas, we were interested in the effect of wetland connectivity on gene flow and continued the analysis without the water land use class. Therefore only five land use classes (i.e., agriculture, hardwoods, manmade structures, pine, and wetlands) were retained in further analyses. For all MLPE model-testing analyses, the full model had the lowest AICc and was therefore considered the best-fit (Table 5). Wetlands were the land use class that explained the greatest amount of variation in genetic distance for the sparse sampling across a large study area and dense sampling across a small study area datasets, and the second most influential land use class in the sparse sampling across a small study area dataset (Table 6). Furthermore, in all three datasets, wetlands consistently had the same sign of effect (i.e., restricting gene flow), and this variable was optimized at the same geographic scale (1000 m) in the sparse sampling across a large study area and dense sampling across a small study area datasets with the same transformation (inverse ricker; Table 7). In the sparse sampling across a small study area dataset

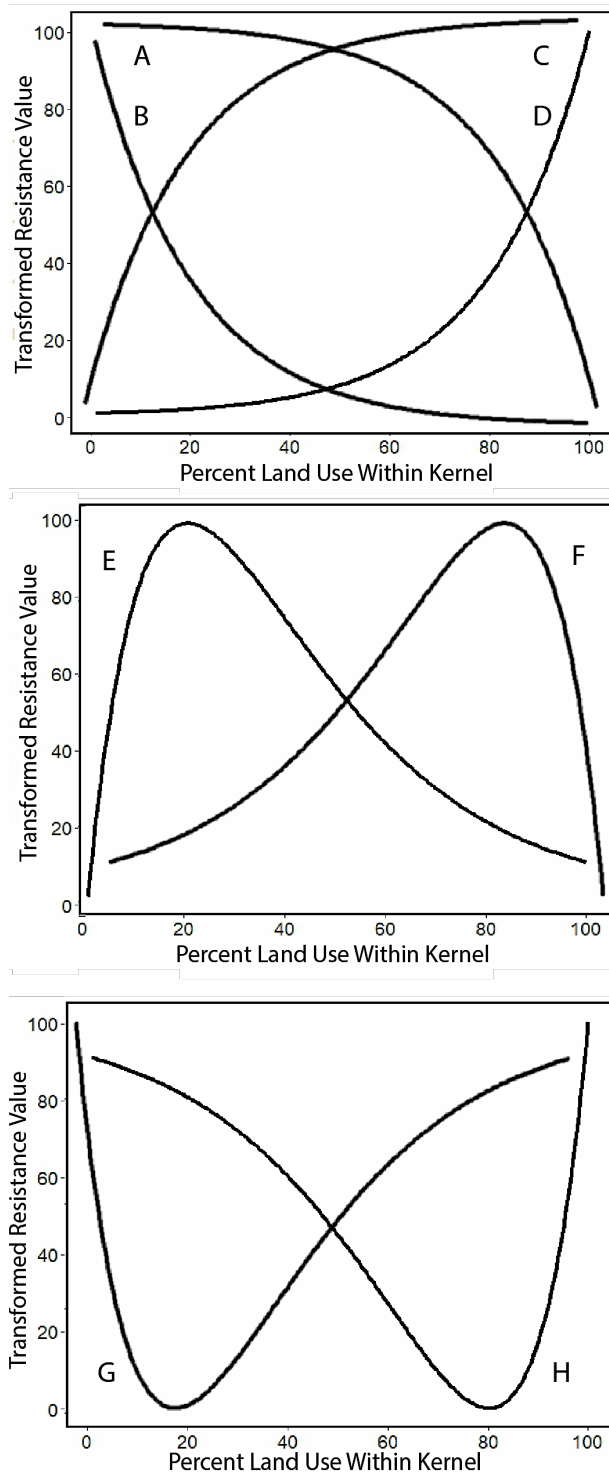


Figure 5. Transformations of original resistance values.

The x-axis shows the original resistance value, which indicates the percent of a given land use class within a moving window. The y-axis represents the new resistance value. The alternative transformations are as follows: A) reverse monomolecular, B) inverse monomolecular, C) monomolecular, D) inverse-reverse monomolecular, E) ricker, F) reverse ricker, G) inverse ricker, and H) inverse-reverse ricker.

Table 5. AICc scores for each multivariate Maximum Likelihood Population Effects model for each set of analyses.

The lowest AICc scores for each category (i.e., sparse/large vs. dense/small vs. sparse/small) are in bold. Land use classes are abbreviated as follows: A=agriculture, H=hardwoods, P=pine, M=manmade structures, and W=wetlands. The effect of geographic distance, calculated using random-walk distance across a homogenous landscape, is represented by GD.

Model name	Variables included	Sparse/large	Dense/small	Sparse/small
Full model	GD, A, H, P, M, W	29467	17565	11517
Isolation by distance	GD	30089	18814	11693
Moderate habitat	GD, A, P	29879	17885	11769
Modified habitat	GD, A, M	29870	17868	11671
Forest cover	GD, P, H, W	29914	17935	11687
Agriculture only	GD, A	29952	18046	11739
Manmade structures only	GD, M	29991	17963	11708
Pine only	GD, P	30025	17966	11702
Hardwoods only	GD, H	29954	18015	11750
Wetlands only	GD, W	29893	17847	11728

Pine Distribution Across Study Areas



Figure 6. Distribution of the “pine” land use class in Holly Springs National Forest.

Map illustrating the distribution of the “pine” land use class across large and small study areas in Holly Springs National Forest.

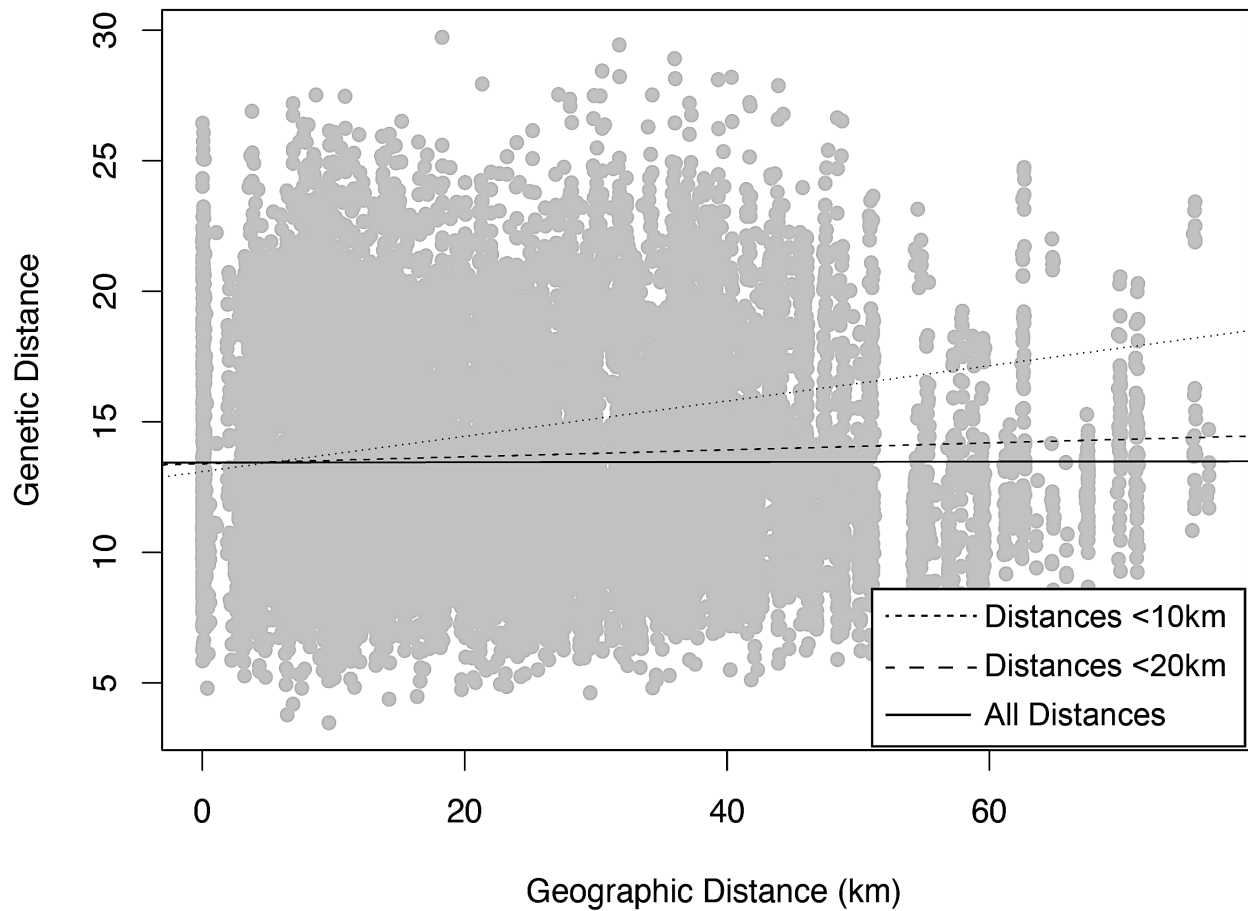


Figure 7. Nested regressions of geographic distance against genetic distance.

Nested regressions of geographic distance against genetic distance. Genetic distance was calculated by first conducting a principle components analysis (PCA) of individual microsatellite genotypes, then calculating pairwise Euclidean distance between PCA coordinates of individuals. Individuals less than 10 km apart: slope=0.07, adjusted $R^2=0.002$ $p=0.004$, individuals less than 20 km apart: slope=0.01, adjusted $R^2=-0.00005$, $p=0.78$, all pairs of individuals: slope=0.0006, adjusted $R^2=0.0001$, $p=0.17$.

Semi-Variogram

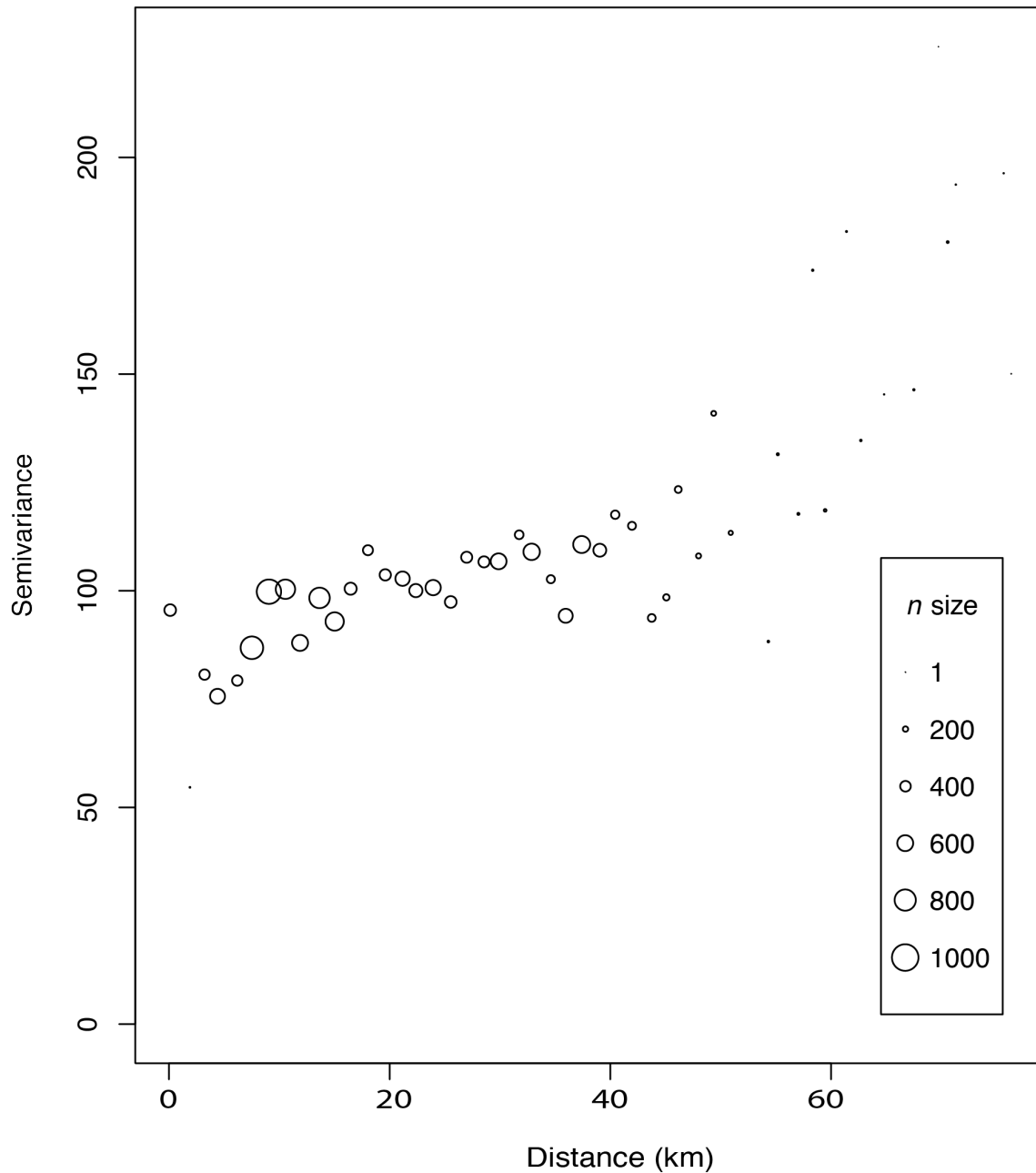


Figure 8. Semivariogram created using pairwise genetic distances and geographic Euclidean distances using 52 distance classes with a distance interval of 1.5 km.

The plateau at approximately 20 km indicates that this is the spatial scale over which spatial autocorrelation is the strongest. N size denotes the number of pairwise comparisons within the given distance class.

wetlands were optimized at a scale of 250 meters with an inverse-reverse monomolecular transformation. These outcomes indicate that for the sparse sampling across a large study area and dense sampling across a small study area datasets, when the amount of wetlands within 1000 m of a given location (i.e., the percent of pixels within 1000 m that are classified as wetlands) is at approximately 20%, gene flow is less restricted. However, as the amount of wetlands within 1000 m approaches 100%, there is a greater restriction of gene flow among *P. mississippi* populations. For the sparse sampling across a small study area dataset, when the amount of wetlands within 250 m is low, gene flow is highly restricted. However, as the surrounding area approaches 100% wetlands, gene flow is less restricted.

The four remaining land use classes were ranked differently among the three datasets (Table 6). Manmade structures were the second most influential land use class in the sparse sampling in a small study area dataset, however in the dense sampling in a small study area dataset they were third, and in sparse sampling in a large study area dataset they were the fourth. The sign and transformation of effect of manmade structures was consistent within the sparse sampling across a large study area and dense sampling across a small study area datasets (inverse ricker). The sparse sampling across a large study area dataset was optimized linearly (Table 7). In all three datasets, the presence of manmade structures correlated with facilitation of gene flow, with low percentages of manmade structures in the nearby area correlating with a small increase in gene flow. As the percent of manmade structures approached 100%, gene flow increased. The scale at which this occurred differed between datasets, with gene flow impacted by manmade structures up to 500 m away in the dense sampling across a small study area dataset and the

Table 6. Rank of effect of landscape variables in best-fit maximum likelihood population effects models for each dataset.

Landscape variables are abbreviated as in Table 5. The number of individuals included in each dataset is indicated by N. Model coefficients, or relative contribution of each landscape variable to genetic distance between individuals, are listed next to each landscape variable.

	Sparse/large dataset N=103		Dense/small dataset N=89		Sparse/small dataset N=89	
Closest distance between sampling locations	7 km		3 km		7 km	
Size of sampling area	630 km ²		256 km ²		256 km ²	
Landscape variable	W	0.91	W	0.81	P	0.84
rank of effect	A	0.69	P	0.68	M	-0.74
and model coefficients	H	0.65	M	-0.60	W	0.67
	M	-0.43	A	0.58	H	-0.41
	P	-0.25	H	0.23	A	-0.09

Table 7. Comparison of the final optimized transformation, scale, and sign of effect for each land use class in all three datasets.

Datasets were developed with either sparse sampling in a large study area (dark grey; >7 km apart), dense sampling in a small study area (light grey; >3 km apart), and sparse sampling in a small study area (white; > 7 km apart, subset area). A positive sign of effect indicates the land use class correlates with a restriction of gene flow, and a negative sign of effect indicates a correlation with facilitation of gene flow.

Land use class	Dataset	Transformation	Scale (m)	Sign
Agriculture	Sparse/large	Inverse-Reverse Ricker	500	+
	Dense/small	Inverse Ricker	500	+
	Sparse/small	Inverse-Reverse Ricker	1000	-
Hardwoods	Sparse/large	Inverse Ricker	500	+
	Dense/small	Reverse Monomolecular	100	+
	Sparse/small	Inverse-Reverse Ricker	750	+
Pine	Sparse/large	Inverse-Reverse Ricker	750	-
	Dense/small	Inverse Ricker	250	+
	Sparse/small	Monomolecular	1000	+
Manmade structures	Sparse/large	Inverse Ricker	250	-
	Dense/small	Inverse Ricker	500	-
	Sparse/small	Linear	500	-
Wetlands	Sparse/large	Inverse Ricker	1000	+
	Dense/small	Inverse Ricker	1000	+
	Sparse/small	Inverse-Reverse Monomolecular	250	+

sparse sampling across a small study area dataset, but only 250 m away in the sparse sampling across a large study area dataset.

Hardwood forests were consistently positively correlated with resistance of gene flow among *P. mississippi* populations (Table 7). However, the relationship between the amount of hardwood forests within the moving window and level of resistance differed across datasets. In the sparse sampling across a large study area and the sparse sampling across a small study area datasets, low percentages of hardwoods within the moving window were correlated with a low level of resistance, and high percentages of hardwoods were correlated with a high amount of resistance. The scales of these effects were different, however, with the sparse sampling across a large study area showing an effect from hardwood forests 500 m away, and the sparse sampling in a small study area showing effects from 1000 m away. In the dense sampling across a small study area dataset, low amounts of hardwoods within a 100 m moving window were correlated with high resistance, whereas high amounts of hardwoods were correlated with lower resistance.

The sign of effect for pine and agriculture varied across the three datasets (Table 7). Pine was correlated with facilitation of gene flow in the sparse sampling across a large study area dataset, however pine forests were correlated with restricted gene flow in both the dense and sparsely sampled datasets within the small study area. In the sparse sampling across a large study area dataset, low amounts of pine within a 750 m moving window correlated with higher gene flow, whereas higher percentages of pine were correlated with less facilitation of gene flow. In the dense sampling across a small study area dataset, the presence of pine was correlated with restricted gene flow. Low percentages of pine within a 250 m moving window were correlated

with low levels of resistance, and higher levels of pine were correlated with high resistance. In the sparse sampling across a small study area dataset, low percentages of pine within a 1000 m moving window were correlated with lower resistance. As the amount of pine increased, resistance increased. Agricultural areas were correlated with resistance of gene flow in the sparse sampling across a large study area dataset and dense sampling across a small study area dataset, however this land use class was correlated with facilitation of gene flow in the sparse sampling across a small study area dataset. In the sparse sampling in a large study area dataset, low amounts of agriculture within 500 m correlated with high levels of resistance, and the lowest amount of resistance to gene flow was found when the window of analysis was comprised of approximately 80% agriculture. In the dense sampling across a small study area dataset, approximately 20% of pine within a 500 m moving window correlated with low resistance. As the amount of agriculture increased, the resistance increased as well. In the sparse sampling across a small study area dataset, pine forests within 1000 m correlated with facilitation of gene flow, with the highest facilitation at approximately 20% pine in the surrounding area, and the lowest facilitation of gene flow when pine forests comprised 80% of the surrounding area.

Discussion

Landscape genetics is still a relatively young sub-discipline, and as such, a number of knowledge gaps remain (Richardson et al. 2016). One of these relates to understanding the effects of sampling density and study area size on inferences about environmental predictor variables that impact gene flow. In this paper, we used a comparison of three landscape genetics

models for *P. mississippi*, a low-mobility salamander species to begin to bridge this knowledge gap. One important consideration in our approach is that the use of ecological predictor variables that have been optimized for scale and transformation separately for each dataset potentially leads to final best-fit models with a number of non-identical features. As such, there can be many nuanced differences among models, some of which may not reflect the specific impacts of sampling density and scale. Accordingly, our assessment of similarities in outcomes generated by the three datasets was necessarily focused at a relatively coarse level (i.e., primarily, which land use classes were included in each best-fit landscape genetic model, their rank ordering of importance, and their overall role in facilitating versus limiting gene flow). Due to the structure of our nested study design, these more substantive differences are likely to be a result of study area or sampling density. Below we discuss the similarities and differences among the sparse sampling across a large study area, sparse sampling across a small study area, and dense sampling across a small study area within the framework of our original research questions (see Introduction). We then close by considering the limitations of our optimized empirical design, and the broader implications of the present study for landscape genetics study design.

The coarsest level of comparison, an examination of the three best-fit MLPE models showed the full model, containing all five land use classes, was the best model in all cases. Wetlands were the most influential land use class in the sparse sampling across a large area and dense sampling across a small area datasets, and the second most influential land use class in the sparse sampling across a small study area dataset. In all three cases there was a correlation between the presence of wetlands and resistance to gene flow among salamander populations. Another notable similarity between all three datasets was consistent correlation between

hardwood forests and resistance to gene flow. Given that *P. mississippi* individuals often reside in bottomland hardwood forests and wetland areas (Petranka 1998), it may seem counterintuitive that high quality habitats such as wetlands and hardwood forests are associated with resistance to gene flow. However, on the basis of empirical (Keely et al. 2016) and simulation (Keely et al. 2017) studies, some researchers have reported that high quality habitat can lead to decreased gene flow among populations, presumably because individuals choose to stay in areas with preferred resources. The final resemblance across all three datasets was a correlation between manmade structures and increased gene flow. This relationship has been seen in other salamander species. Prunier et al. (2014) found a correlation between roads (a significant component of the “manmade structure” land use class in our analysis) and increased gene flow among alpine newt (*Ichthyosaura alpestris*) populations. Although avoidance of road edges has been documented in many amphibian species, Marsh and Beckman (2005) found no effect of forest roads on the presence of slimy salamanders (*Plethodon glutinosus*, the sister species to *P. mississippi*) supporting the idea that they may move freely across them. Thus, the correlation between facilitation of gene flow and manmade structures may be species specific.

Based on the nested design of our sampling, differences between the datasets sampled densely and the datasets subsampled sparsely across a small study area and the dataset sampled sparsely across a large study area can shed light upon the effect of study area on landscape genetic inferences. For instance, if the best-fit models for sparse and dense sampling across a small study area were similar to each other, but differed from the large study area (only sparsely sampled for purposes of tractability), this would indicate that study area size was important in its own right. In our analyses, we found this effect with respect to pine forests. When comparing

datasets within the small study area to the large study area dataset, we found contrasting directionality (i.e., sign) of the correlations between presence of pine forests and gene flow among *P. mississippi* populations. Specifically, the analyses based on sparse sampling across a large study area detected a correlation between pine and facilitation of gene flow, whereas the analyses based on dense and sparse sampling across a small study area identified this land use class as being associated with restricted gene flow (Table 7). A number of factors may have led to different inferences, including the configuration and variability of land use patches across the landscape. For example, in landscape genetics analyses of the American pika (*Ochotona princeps*) Castillo et al. (2016) found the configuration (patch connectivity) of a given environmental predictor variable impacted the relationship between the variable and gene flow. Other landscape genetics studies have found habitat patch characteristics (e.g., density, cohesion, and correlation length) can drive different outcomes in replicated or nested analyses (Cushman et al. 2012; Cushman et al. 2013; Vergara et al. 2017). When comparing these patch configuration metrics in the small versus large study area, we find that the patch density of pine forests within the small study area is much lower than the pine patch density within the large study area (Table A4, Figure 6). Although patch density calculations can be affected by the size of the area in which they are calculated (McGarigal et al. 2012), the difference in pine patch densities is far greater than the differences found among the patch densities of agricultural areas, hardwood forests, manmade structures, and wetlands in the two study areas. These results support the idea that patches of pine may lead to increased *P. mississippi* dispersal if they are densely positioned across the habitat matrix, however, when patches of pine forest are more fragmented they lead to decreased dispersal and resistance to gene flow.

The final landscape variable, agriculture, followed a pattern that does not clearly support an effect of study area size or sampling density, but may instead be a combination of effects from sampling density and study area size (Figure 4D). Here, we break down the potential effects of both study area size and sampling density on these different inferences. The densely sampled dataset across the small study area resulted in inferences similar to those found using a large study area (correlation with resistance to gene flow, Table 7), however the dataset that represented sparse sampling across a small study area did not, instead correlating with facilitation of gene flow. While our results indicated the difference between inferences generated in the large and small study areas with respect to pine were driven by patch configuration, the patch density, patch cohesion, and correlation length of agricultural patches in the large and small study areas were relatively similar. However, our agricultural land use class includes a wide array of crops, including soy beans, cotton, sweet potatoes, and corn. Thus, it is possible that the agricultural areas found within the small study area are not a representative subset of the large study area, and the relationship between gene flow among *P. mississippi* populations and agriculture is dependent upon crop type.

We also found differences in the inferences generated by datasets sampled sparsely and densely across a small study area. Agricultural areas correlated with resistance to gene flow when sampled densely across a small study area, however they were correlated with facilitation of gene flow when subsampled sparsely across the same area. Based on the nested design of our sampling and analysis, differences between sparse sampling and dense sampling would indicate the relationship between long distance dispersers and environmental variables may be different than that of short distances dispersers and that environmental variable. In many species, dispersal

can be divided into two categories: a movement that is the result of several small, “routine” movements, versus a singular “special” movement (Van Dyck and Baguette 2005). For example, in an experimental analysis of the ringed salamander, *Ambystoma annulatum*, Ousterhout and Semlitsch (2018) found individuals could be identified as “resident” salamanders that dispersed small distances, and “dispersing” salamanders that dispersed greater distances. In that study, the movement of residents was not impacted by habitat type, whereas movement of dispersing salamanders was, with dispersing individuals moving farther through grasslands than hardwood forests. In a landscape genetics analysis, this differentiation between the behavior of short versus long distance dispersers would likely manifest as different inferences when using dense versus sparse sampling. While our results did not match a pattern that supported a clear effect of sampling density on landscape genetics inferences (i.e. both sparsely sampled datasets resulted in similar inferences that were different from the densely sampled dataset), we did find contrasting directionality (i.e., sign) of the correlations between presence of agricultural areas and gene flow among *P. mississippi* populations. The potential for scale-dependent relationships between movement and environmental heterogeneity of the intervening landscape has also been reported in other studies of dispersal behavior in mammal and amphibian species. For example, for Iberian lynx (*Lynx pardinus*), agricultural areas cause resistance to movement within an individual’s home range, however when dispersing long distances, agricultural areas and manmade structures caused less resistance (Gastón et al. 2016). In the invasive cane toad (*Rhinella mariana*), this relationship also appears to be affected by whether an individual is at the core versus edge of the species’ range, as individuals at range margins exhibited more long distance, exploratory movement (Gruber et al. 2017). Collectively, these studies show that dispersal behaviors can vary, and a component of that variation is often a different relationship

between environmental variables and dispersal distance. Our findings suggest a there may be a similar relationship between dispersal distance and the presence of the type of agricultural areas found in the small study area, where short distance dispersers did not regularly disperse through agricultural areas, yet long distance dispersers more readily moved through it.

The present study was examined the effects of sampling density and study area size on landscape genetics inferences, via reanalysis and comparison of empirical datasets. To date, landscape genetics methods have largely been explored using simulations (e.g., Cushman et al. 2010; Landguth et al. 2012). While empirical studies have the ability to assess interactive and additive forces (Resasco et al. 2017), their use in testing the effect of different methods has limitations. For example, geographic replicates are never identical, and so as in the present study, multivariate models created using optimized predictor variables (e.g., land use classes) often include predictor variables of different scales and transformations (Castillo et al. 2016; Vergara et al. 2017). The overall size of HSNF also limited the possible combinations of sampling density in our nested design. For example, we did not include dense sampling across a large study area, or more than one subset area. We also did not include a very fine-scale sampling strategy in which geographic spacing among sampling locations matched average individual dispersal distance, because a feasible number of sampling locations would encompass a very small area unlikely to include all five land use classes. Despite these limitations, the use of nested empirical analyses allows for a greater understanding of the effect of landscape genetics methodology that is limited by the complexities of empirical datasets (e.g., non-uniform sampling, previously undefined population structure, missing data, etc).

Implications for sampling design of empirical landscape genetics studies

The nested sampling design used in the present study detected relationships between gene flow and the environment that were unique to study areas as well as to short distance comparisons. Although the relationship between gene flow and wetland areas, manmade structures, and hardwood forests were similar regardless of sampling density and study area size, analyses revealed patch density likely affected the relationships between gene flow and pine forests, and agricultural type may have played a role in the relationships between agricultural areas and gene flow. These findings highlight the strength of potential influence patch configuration may have on the relationship between gene flow and environmental variables. They also support the use of finer categorical resolution when determining the effect of agriculture on salamander dispersal. In order to capture the potential effects of habitat configuration on landscape genetics inferences, we recommend conducting analyses in multiple study locations with different patch configurations. Furthermore, inferences revealed that within the small study area, *P. mississippi* dispersers that disperse long distances move more readily through agricultural areas than those that disperse short distances. These findings are consistent with the idea that the relationship between the environment and dispersal via repeated routine movements (i.e., short distance dispersal) is not the same as for special, long distance dispersal. Furthermore, our results indicate these differences between short and long distance dispersal are different for each land use class. Thus, in order to understand the extent to which short and long distance dispersers are affected by the environment, we recommend the use of a dense sampling scheme followed by reanalysis of these data using “sparse” pairwise comparisons. By carrying out landscape genetics analyses at in multiple areas of diverse patch configurations as well as

multiple sampling densities, researchers may identify correlations between gene flow and the environment that are unique to habitat configurations or short versus long distance dispersal.

CHAPTER 4

ANALYSIS OF CUTANEOUS AND LOCAL SOIL MICROBIOMES IN A TERRESTRIAL SALAMANDER: THE ROLE OF SIBLINGS VERSUS SOIL

Abstract

The mutualistic relationship between amphibians and their cutaneous microbial community can strengthen the host's ability to fight pathogens such as *Batrachochytrium salamandrivorans* (*Bsal*). Manipulation of the amphibian cutaneous microbiome via probiotic inoculation is a promising strategy for mitigation and containment of *Bsal* outbreaks. An understanding of the mechanisms by which amphibians acquire their cutaneous microbiome is pivotal to the development of effective probiotic mixtures. Using microsatellite-based salamander genotypes and 16s rRNA microbiome characterization, we investigated the impact of genetic relationships on salamander cutaneous microbiomes, as well as the relationship between the individual salamander microbiomes and the microbiomes of their immediate environment (i.e., soil). Neither relatedness nor kinship of salamanders was correlated with the composition of their cutaneous microbiomes. Approximately half of salamander microbiomes were similar to soil microbiomes, while the other half clustered separately in ordination space. The unique components of the salamander microbiomes were genetically similar to antifungal operational taxonomic units (OTUs). In a series of exact tests between salamander cutaneous microbiomes

and the microbiomes of their immediate environment, salamanders showed higher, statistically significant abundance of members of the family Methylobacteriaceae, which is known to include several antifungal taxa. However, several families with antifungal properties were found to be more abundant in soil microbiomes, suggesting the mere presence of potentially beneficial bacteria does not necessarily lead to their incorporation into amphibian cutaneous microbiomes. Knowledge of this relationship could be used to in the use and development of environmental probiotic mixtures.

Introduction

The emerging fungal pathogen *Batrachochytrium salamandrivorans* (*Bsal*) has been identified as the cause of mass fatalities in salamander populations throughout Europe (Marten et al. 2014; Spitzen-Van Der Sluijs et al. 2016). *Bsal* has not yet been introduced into North America, however, if the fungus reaches the continent, disease risk assessments have projected significant decreases in salamander biodiversity within the Pacific, southern Appalachian, and mid-Atlantic regions (Richgels et al. 2016). Accordingly, researchers, managers, and policy makers throughout North America have focused efforts on preparing for the potential outbreak of *Bsal* by developing response plans that include mitigation and containment measures (Grant et al. 2015; *Bsal* Task Force 2019). An important facet of these plans is increasing the resilience of salamander populations to *Bsal* infection by understanding and potentially manipulating the mutualistic relationship between amphibians and their native cutaneous microbiome (Becker and Harris 2010; Woodhams et al. 2014). Although some of the microbiome includes fungal and

microeukaryotic species that may prove to be beneficial (Kueneman et al. 2016), the primary focus of amphibian cutaneous microbiome research (and the focus of this study) is on beneficial bacterial taxa. Bacterial isolates from amphibian skin that inhibit the growth of fungi have been identified (Harris et al. 2009; Becker et al. 2009), leading to a collaborative database of amphibian skin bacterial isolates that exhibit antifungal properties in laboratory trials (Woodhams et al. 2015). Currently, understanding the mechanisms by which individuals acquire these antifungal bacteria is ongoing.

As in any community, microbiome species composition and acquisition is often the product of responses to biotic and abiotic factors, such as species invasion, species loss, and environmental variability (Prosser et al. 2007). To predict the likelihood of individual salamanders incorporating probiotic bacterial species into their cutaneous microbiome, managers first need to understand how cutaneous microbial communities interact with environmental microbial communities. Experimental manipulation has shown amphibians rely on their environment as a species reservoir that can increase the diversity of their cutaneous microbiome. For example, salamanders housed in sterile laboratory conditions exhibit lower microbial species diversity than individuals housed with their native soil (Loudon et al. 2014) and individual salamanders housed in soil inoculated with a proposed probiotic *Janthinobacterium lividum* incorporate the bacterium into their cutaneous microbiomes (Muletz et al. 2012). However, in these experiments, individuals' existing microbiomes were removed or disturbed, either through the use of antibiotics (Loudon et al. 2014) or hydrogen peroxide (Muletz et al. 2012). A significant knowledge gap currently exists regarding the relationship between intact cutaneous microbiomes and environmental microbiomes. Thus, to predict the likelihood of individual

salamanders incorporating probiotic bacterial species into their cutaneous microbiome, managers need to first understand how cutaneous microbial communities interact with environmental microbial communities.

The composition of amphibian cutaneous microbiomes has been attributed to both environmental inocula and host selection pressures (e.g., via peptide secretions; Walke et al. 2014). Researchers have found the microbiomes of individuals within the same location (i.e., the same habitat) are more similar to one another than to the microbiomes of individuals from different locations, regardless of the salamander species (Muletz-Wolz et al. 2018) or even genus (Bird et al. 2018) to which individuals belong. This suggests the effect of the local environment is more important than the selective pressures unique to a given species. However, when comparing the microbiomes of individuals from different orders (e.g., frogs versus salamanders) the difference in selection pressures is more important than habitat (Ellison et al. 2018). These studies provide support for the influence of broad host-specific selective pressures on cutaneous microbiomes, however they have not investigated the effect of fine-scale host genetic differentiation on cutaneous microbiome composition, such as the potential variability among distinct populations of a focal species (Becker et al. 2017). A significant positive relationship between the kinship of two individuals and the similarity of their cutaneous microbiomes could indicate a type of natal influence, wherein the microbiome of an individual's natal environment impacts the structure of their cutaneous microbiome later in life. The effect of kinship has been found to have a significant impact on the fecal microbial similarity in tortoises (Yuan et al. 2015). If individual relatedness is significantly positively correlated with cutaneous microbiome similarity, the use of microbiome manipulation (i.e. probiotics) as a response to disease spread

may be differentially effective across salamander populations, as effectiveness may depend on the birthplace of an individual.

In the present study, we investigated the impact of relatedness on salamander cutaneous microbiomes, as well as the relationship between the individual salamander microbiomes and the microbiomes of their local environment (i.e., soil). The Mississippi slimy salamander (*Plethodon mississippi*) is a low mobility species (Wells and Wells 1976) that resides under rotting logs in the bottomland hardwood forests of Mississippi and Alabama. *P. mississippi* is within the same species complex as *P. glutinosus*, which has been shown to exhibit transient skin invasion when experimentally infected with *Bsal*, meaning member of this species complex may serve as a reservoir or carrier of the fungus in the event of a North American invasion (Martel et al. 2014). Using *P. mississippi* relatedness and the individual salamander's cutaneous microbiomes, we asked whether genetic relatedness between individuals predicts the similarity of their cutaneous microbiomes. We also compared salamander cutaneous microbiomes to the microbiomes of their local soil environment to determine whether microbial taxa in cutaneous microbiomes are differentially abundant in the salamander's immediate environment.

Methods

Sampling

In April 2017, we sampled 70 *Plethodon mississippi* individuals from 14 locations spaced approximately 3 km apart in Holly Springs National Forest in northeastern Mississippi. Upon capture, each individual was placed in a clean plastic container and rinsed with distilled water to remove soil and transient bacteria. Each salamander was then swabbed 20 times each on its ventral and dorsal sides with a sterile polyester-tipped applicator. The applicator was placed in a Qiagen DNeasy Powersoil Kit (Valencia, CA, USA) collection tube and frozen at -20°C for storage. A tail tip was also taken from each individual and placed in ethanol for DNA extraction and genotyping. A sample of soil from the immediate location where the salamander was found was also placed in a separate Qiagen DNeasy Powersoil Kit collection tube and frozen at -20°C for storage. Sampling was approved by the University of Mississippi IACUC (#15-020), the Mississippi Department of Fish and Wildlife (Permit #0324164), and the USDA Forest Service.

DNA Extraction, Bacterial Sequencing, and Salamander Genotyping

Bacterial genomic DNA was extracted from skin and soil samples with the Qiagen DNeasy Powersoil Kit using manufacturer-recommended protocol. The V4 region of bacterial 16s rRNA gene was amplified using a dual-index barcoding approach (Kozich et al. 2013). Polymerase chain reactions (PCRs) contained 1 µL genomic DNA and 17 µL AccuPrime Pfx Supermix (Invitrogen, Grand Island, NY). PCR was performed for 30 cycles of 95°C for 20 s, 55°C for 15 s, and 72°C for 2 min with a preliminary 95°C denaturation for 2 min and a final 72°C elongation for 10 min. Products were standardized with SequalPrep normalization plates (Life Technologies, Grand Island, NY), pooled, and sequenced at the Molecular and Genomics Core Facility at the University of Mississippi Medical Center using an Illumina MiSeq platform.

Raw FASTA sequences were processed using the mothur v. 1.35.1 (Schloss et al. 2009) pipeline following the protocol detailed by Jackson et al. (2015) to remove sequencing errors, chimeras, and sequences with ambiguities. Remaining sequences were aligned using the SILVA version 132 database (Quast et al. 2013), and classified using the RDP (release 11) database (Wang et al. 2007). Highly similar bacterial sequences (>97% similarity) were pooled together to form operational taxonomic units (OTUs).

Salamander DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit (Valencia CA, USA) using manufacturer-recommended protocol. Individuals were genotyped at eight microsatellite loci described by Spatola et al. (2013; see Appendix and Tables A1, A2 and A3 for PCR conditions and allele-calling approaches). The resulting dataset of multilocus microsatellite genotypes was tested for adherence to basic molecular marker inheritance patterns by grouping all individuals into a single putative population. The R (R Core Team 2019) package “PopGenReport” (Adamack and Gruber 2014) was used to test for deviations from Hardy-Weinberg Equilibrium and null alleles, and to calculate percent missing data, number of alleles per locus, and mean allelic richness. We tested for departures from linkage equilibrium using MICRO-CHECKER v 2.2.3 (van Oosterhout et al. 2004).

Statistical Analysis

Using the most common sequence as a representative sequence for each OTU, an approximately-maximum likelihood tree of OTUs was created with a generalized time reversible model of DNA sequence evolution in the program FastTree 2.1.11 (Price et al. 2010). The

resulting tree was combined with mothur outputs in the R package “phyloseq” (McMurdie and Holmes 2013). To determine the presence and composition of a core microbiome among salamanders, “phyloseq” was used to identify only those OTUs that were present in at least 80% of salamander samples. To describe differences in microbial community diversity among samples, a subset of OTUs present in at least 25% of all samples (i.e., salamander and soil) was used to calculate pairwise weighted UniFrac (Lozupone and Knight 2005) distance among all samples and the UniFrac distances were plotted using Principle Coordinates Analysis (PCoA). The effects of microbiome origin (i.e. either soil or salamander) and site on weighted UniFrac distances were tested using a PERMANOVA with 999 permutations. Each OTU in the subset of OTUs that were present in >25% of all samples was subjected to a custom BLAST against the Amphibian Skin Antifungal Isolates 16s rRNA Database (Woodhams et al. 2015). OTUs were classified as “inhibitory” if their identity to an isolate labeled inhibitory was >97%. An asymptotic Wilcoxon-Mann-Whitney test was then used to test for differential distribution of “inhibitory” and “non-inhibitory” OTUs along axis 1 of the PCoA. To test for differential abundance of bacterial families in paired salamander and soil samples, (i.e., when comparing salamander microbiomes to the microbiomes of soil collected in their immediate environment), our dataset was first agglomerated to the family taxonomic level. The R package “edgeR” (Robinson et al. 2010) was used to conduct a series of exact tests between paired microbiomes while accounting for the differences in average library size between salamander and soil samples using variance stabilization.

Pairwise relatedness (r) between individual salamanders was estimated using maximum likelihood in the program ML-Relate (Kalinowski et al. 2006). The inverse of relatedness ($1-r$)

was calculated to generate a measure of dissimilarity between individuals. The resulting dissimilarity matrix was tested for correlation with pairwise UniFrac distance between salamander cutaneous microbiomes using a mantel (Mantel 1967) test in the “vegan” (Oksanen et al. 2019) package in R with 9999 permutations. Salamander cutaneous samples clustered in two distinct groups along axis 1 of the PCoA of UniFrac distances described above. Individuals were divided into clusters according to their location along PCoA axis 1: cluster A consisted of individuals that grouped closely with soil microbiomes and had axis 1 scores <0.11 , whereas cluster B consisted of individuals whose microbiome grouped separately from soil microbiomes and had axis 1 scores >0.11 . A Welch two-sample t-test using base R functions was conducted to test for a significant difference between the mean relatedness of individuals within PCoA clusters.

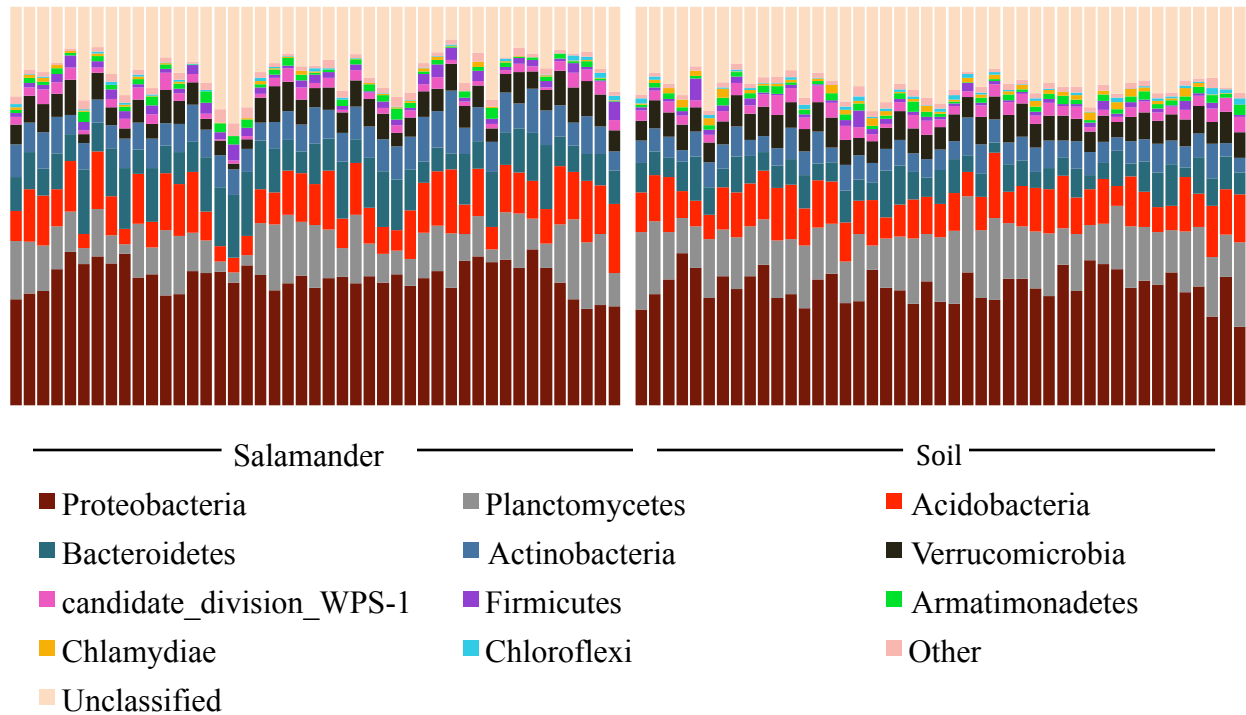
Results

Of the 70 soil samples, all 70 showed successful amplification at the 16s rRNA V4 region and were sent to the Molecular and Genomics Core Facility at the University of Mississippi Medical Center. Of the 70 salamander cutaneous swabs, 57 were successfully amplified at the V4 region. Following initial analysis to remove sequencing errors, chimeras, sequences with ambiguities, and non-bacterial (chloroplast, mitochondrial, Eukarya, and Archaea) sequences, 45 salamander swab samples and 45 corresponding soil samples were retained for further analysis on the basis of rarefaction curves that approached level. The 90 microbiome samples retained contained 1,283,727 sequences, of which 111,212 were unique and clustered into 27,906 OTUs.

Both salamander cutaneous and soil microbiomes contained a large percentage of Proteobacteria, as well as Planctomycetes, Acidiobacteria, and Bacteroidetes (Figure 9A). OTUs within the phylum Proteobacteria were mostly Alphaproteobacteria (Figure 9B).

The salamander core microbiome consisted of 12 OTUs that were present in at least >80% of salamander swab samples (Table A5). A total of 276 OTUs were present in at least 25% of both salamander and soil samples and were used for UniFrac calculation and paired exact tests. Five of the 12 “core” OTUs were identified as inhibitory using the Amphibian Skin Antifungal Isolates 16s rRNA Database. Similarly, 27 of the 276 OTUs used to describe microbial community diversity in UniFrac calculations were identified as inhibitory. UniFrac distances did not differ statistically by site ($p=0.273$), however they differed significantly by microbiome origin (i.e. salamander or soil, $p<0.001$). PCoA axis 1 explained 68.6% of variance found in salamander and soil in microbiome diversity, with 26 salamander samples clustering with soil samples along axis 1, and 19 clustering separately (Figure 10A). Several OTUs also clustered with the differentiated salamander cutaneous samples along axis 1 (Table A6). Of the 14 OTUs with axis 1 scores > 0.2 , three OTUs were identified as inhibitory by >97% identity with OTUs listed as “inhibitory” in the Woodhams et al. (2015) antifungal database. An asymptotic Wilcoxon-Mann-Whitney test showed inhibitory and non-inhibitory OTUs were differentially distributed along axis 1 ($p<0.0001$) with inhibitory OTUs closer to cluster B (non-inhibitory axis 1-inhibitory axis 1 = -0.0388). Exact tests of salamander-soil pairs showed 66 bacterial families were differentially abundant (log-fold change $> |2|$, $p<0.001$) in salamanders versus their local soil environment (Figure 11). Five families were more abundant in salamander cutaneous microbiomes, and 61 families were more abundant in soil microbiomes.

A) Relative Abundance of Bacterial Phyla



B) Relative Abundance of Proteobacteria Subphyla

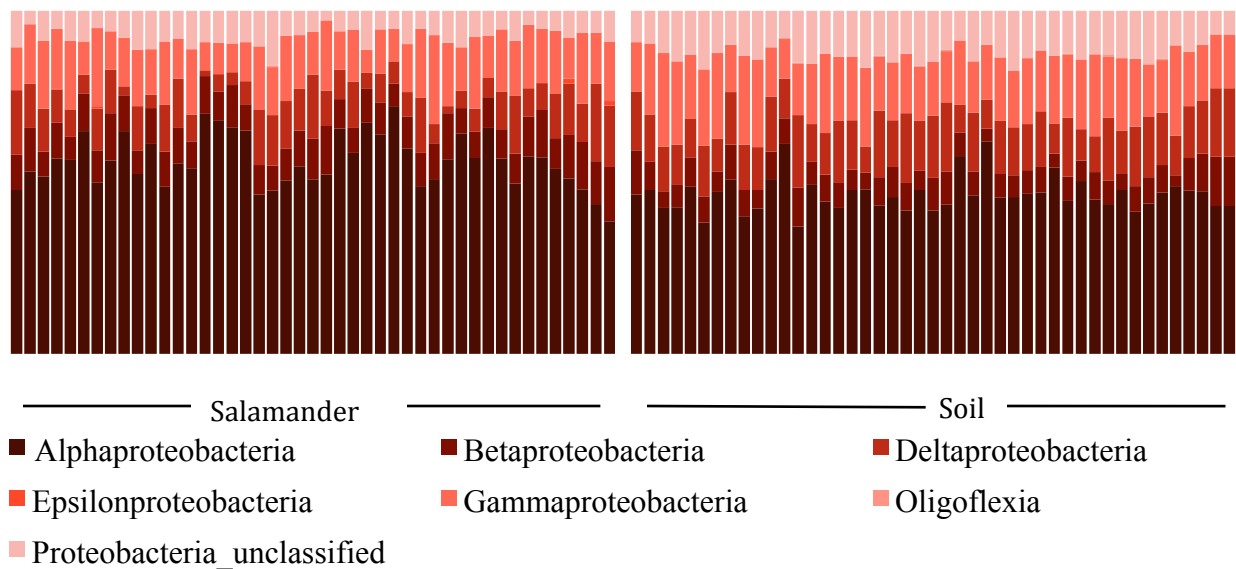
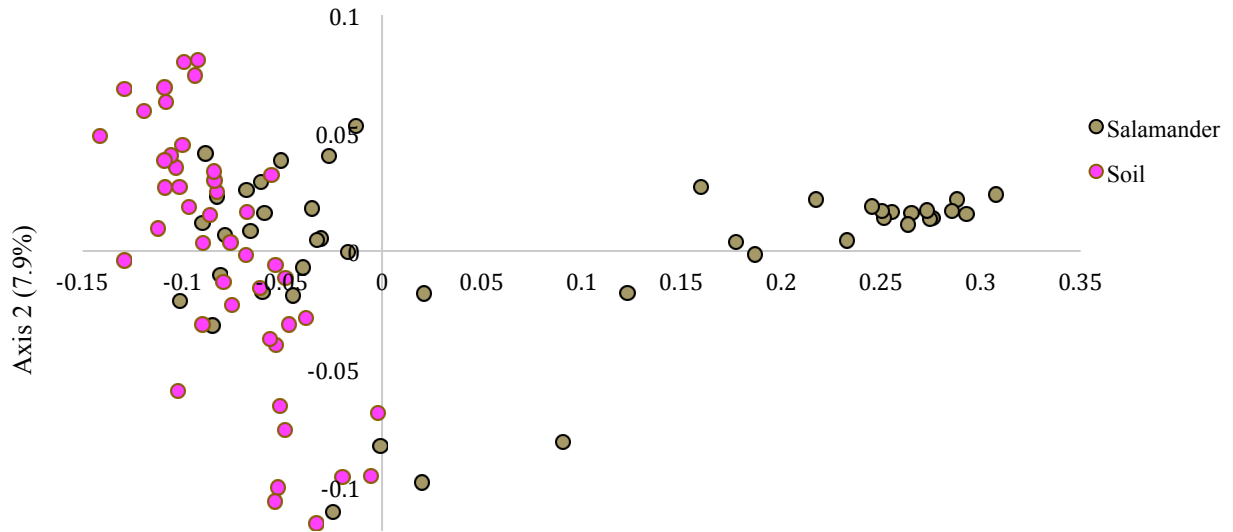


Figure 9. Relative abundance by sample of a) common phyla and b) subphyla within Proteobacteria.

Samples are sorted into salamander swab samples (left) and soil samples (right). Within the salamander and soil groupings, samples are arranged by site.

A) Principal Coordinates Analysis of Salamander and Soil Microbiome UNIFRAC Distances



B) OTU Scores

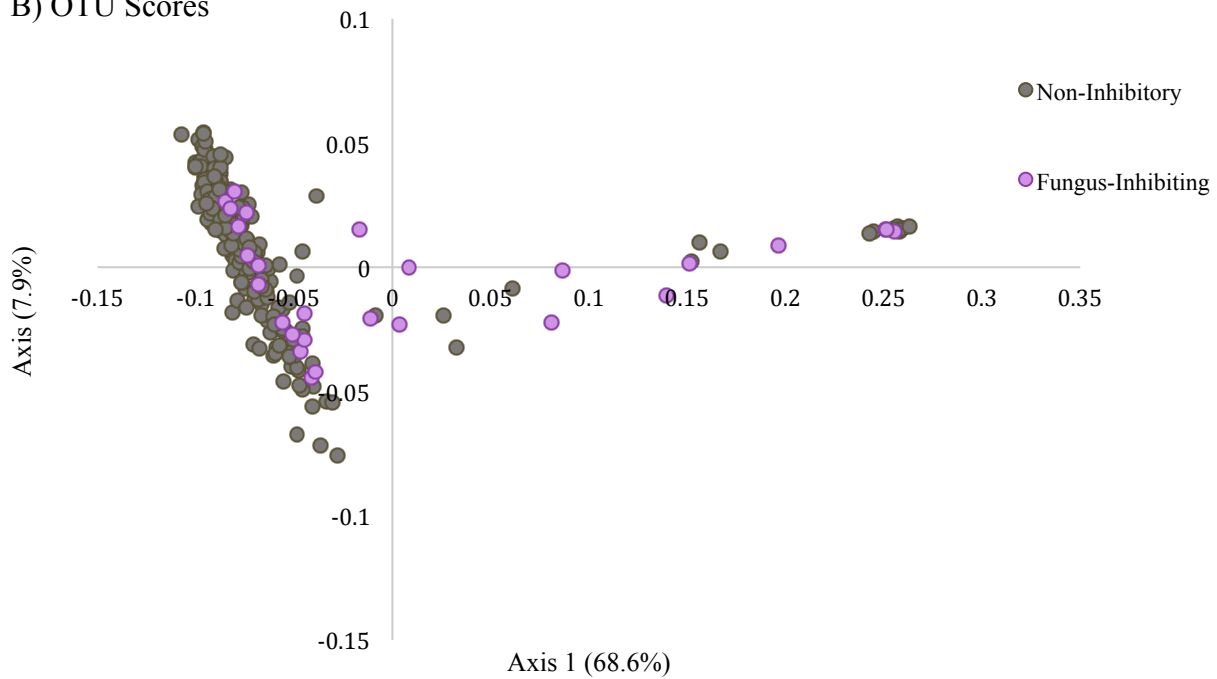


Figure 10. Principal Coordinates Analysis (PCoA) of pairwise UNIFRAC distances between all microbiome samples.

A PCoA of weighted UNIFRAC distances between all microbiome samples shows clustering of soil and half of salamander microbiomes, with differentiation of the remaining salamander samples along Axis 1 (68.6% of variation). B) Species scores from the above PCoA classified as either “fungus-inhibiting” or “non-inhibitory” by their >97% identity match to bacterial isolates known to inhibit *Batrachochytrium dendrobatidis* (*Bd*) (Woodhams et al. 2015).

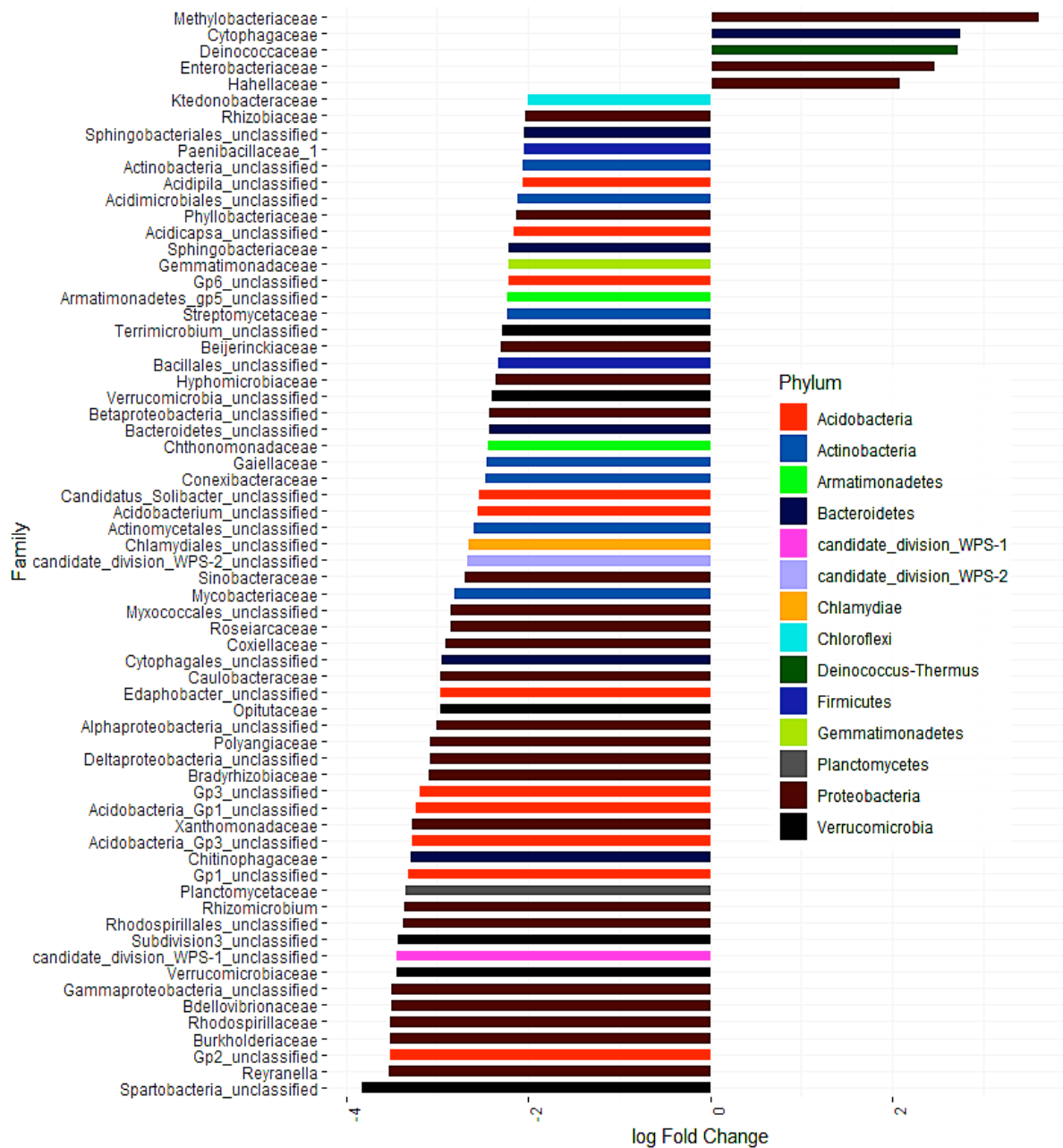


Figure 11. Bacterial families with significant differential abundance between paired samples of salamander cutaneous microbiomes and the microbiomes of their immediate environment.

An exact binomial test that accounts for very different total species counts was used to test for differential abundance between soil and salamander microbiomes. Families with a negative log-fold change show greater abundance in soil microbiomes when compared to the microbiomes of salamanders caught in that immediate location. Families with a positive log-fold change show greater abundance in salamander microbiomes than the microbiomes of the salamander's immediate environment. For all listed families, $p < 0.001$

Multi-locus genotypes were successfully generated for all 70 salamanders, however the only genotypes included in this study were those of the 45 individuals with cutaneous microbiome data. Within this dataset, total missing genotypic data was 2.5%. The number of alleles per locus ranged from 3-19, with a mean of 9.8 alleles per locus. Tests for deviation from Hardy-Weinberg Equilibrium identified homozygote excess at locus *43M*, suggesting the possibility of a null allele. The locus was retained for further analysis, however *43M* was identified as locus with null alleles when calculating pairwise relatedness and kinship in ML-Relate. Of the possible 990 pairwise combinations of salamanders, there were 44 pairs of half siblings, 3 pairs of full siblings, and 11 parent-offspring pairings. A Mantel test comparing inverse relatedness to UniFrac distance (between salamanders only) was not significant (Mantel $r = -0.04913$, $p=0.9438$). Mantel tests comparing inverse relatedness to UniFrac distance in cluster A (PCoA axis 1 scores <0.11) and cluster B (PCoA axis 1 scores >0.11) were also not significant (cluster A: Mantel $r = -0.04816$, $p=0.7864$; cluster B: mantel $r = -0.0983$, $p=0.9032$). Cluster A contained 10 parent-offspring pairs, 2 pairs of full siblings, and 22 pairs of half siblings. Cluster B contained 1 parent-offspring pair, 1 pair of full siblings, and 22 pairs of half siblings. The mean UniFrac distances between parent-offspring pairs, full siblings, half siblings, and unrelated pairs were not statistically different (Figure 12, Kruskal-Wallis $p=0.31$).

Discussion

We tested for correlation between genetic relatedness of individuals and the similarity of their microbiomes, which, if significant, would indicate individual cutaneous microbiomes are

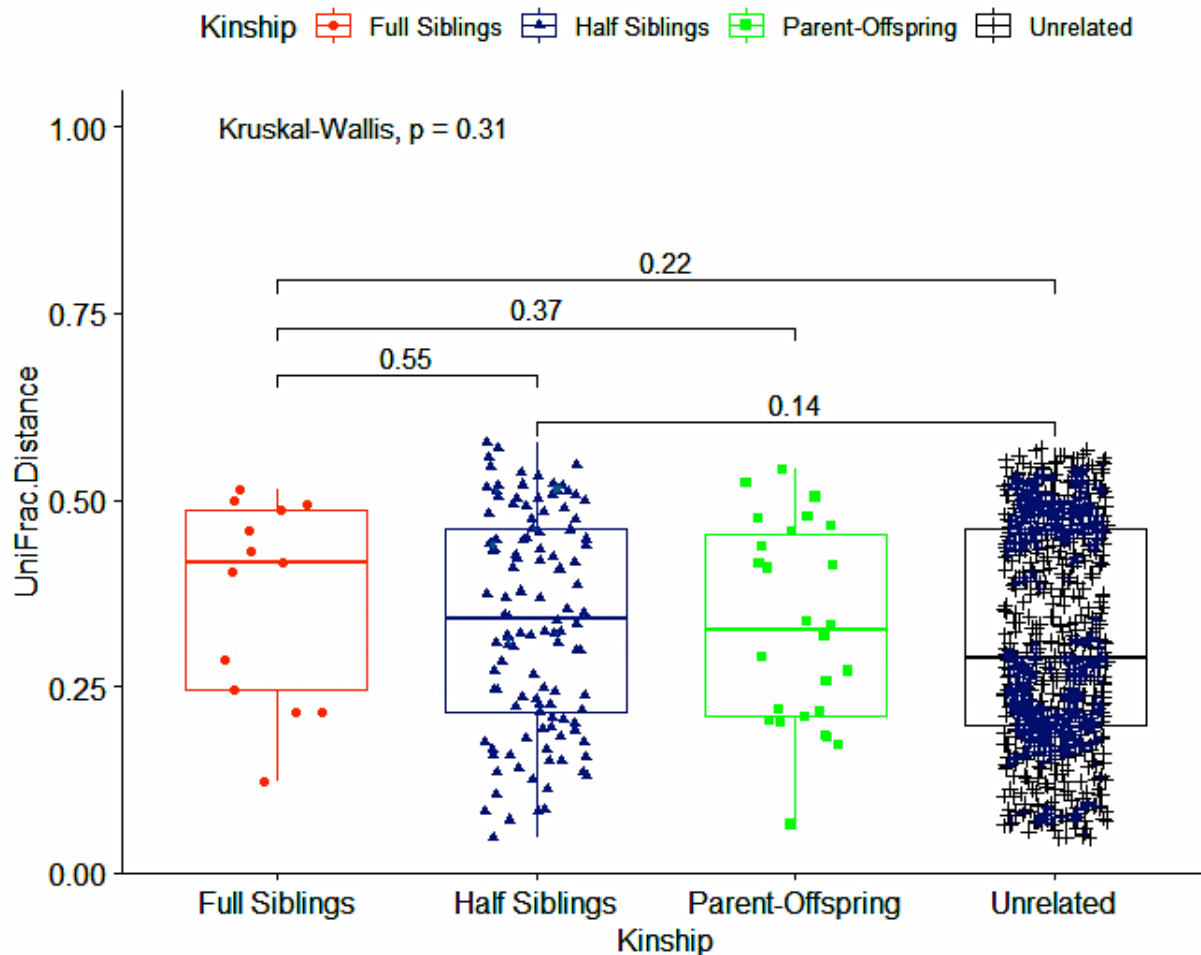


Figure 12. UniFrac distances by kinship category.

Mean UniFrac distances between individuals identified as full siblings, half siblings, parent-offspring pairs, and unrelated pairs were not statistically different ($p=0.31$). Box plots illustrate the 75th percentile (top), 50th percentile (median, middle line), and 25th percentile (bottom) of pairwise UniFrac distances. Dots, triangles, squares, and crosses represent pairwise UniFrac distances between full siblings, half siblings, parent-offspring pairs, and unrelated individuals respectively.

impacted by their natal environment. However, our results show a lack of correlation between relatedness or kinship and cutaneous microbiome composition (Figure 12). To date, the diversity of amphibian cutaneous microbiomes has not been shown to correlate with relatedness or kinship. Within a different microbial environment, Griffiths et al. (2018) found distance between tadpole mouthpart microbiomes significantly correlated with genetic distance between Phofung river frog (*Amietia hymenopus*) hosts. However, frog population genetic patterns were strongly shaped by a river network, where populations that were adjacent along the river were more genetically similar than those that were not. Yet the distances among mouthpart microbiomes did not follow this same pattern. Their results illustrate that it is often difficult to determine whether the similarities between microbiomes of individuals in a shared environment are due to host relatedness or the environment itself. Because we did not see a relationship between relatedness or kinship and the composition of cutaneous microbiomes, we propose *P. mississippi* cutaneous microbiomes are largely a contemporary product of their environment, as opposed to a product of their natal environment.

In our investigation of salamander and soil microbiome diversity, we found 26 of 45 salamander microbiomes were similar to soil microbiomes (Figure 10A). The remaining salamander microbiomes clustered separately along an axis of ordination that explained a large portion of microbial variance (68.6%). Of the OTUs that clustered with this group, three are antifungal, and an additional four are within families with known antifungal species. However, while the salamanders that clustered separately from soil have an increased number of antifungal OTUs when compared to the other cluster, all salamander samples contained some antifungal OTUs. Of the OTUs present in at least 80% of salamander samples (i.e. the core

salamander microbiome), 42% were antifungal. Furthermore, paired salamander-soil exact tests show a significant log-fold increase in the abundance of OTUs within the Methylobacteriaceae family, which contains a considerable number of antifungal species. From this information, it is clear that antifungal OTUs are commonly part of the salamander cutaneous microbiome.

However, several of the bacterial families that were found to be more common in the soil than on the salamander also contain antifungal OTUs, suggesting the mere presence of potentially beneficial bacteria in the environment does not guarantee their incorporation into amphibian cutaneous microbiomes.

It is important to note that this study offers only a snapshot of salamander microbiomes and the microbiomes of their environment, which may be influenced by biotic or abiotic factors. Clinical trials have shown decreases in antifungal OTUs on salamander skin due to increased temperatures (Muletz-Wolz et al. 2019), and soil microbiomes can be influenced by a number of factors, including micropredators, carbon availability, and environmental pH (Fierer et al. 2017). Furthermore, because these are wild-caught individuals, we lack knowledge of the history of the salamanders infections, wounds, or interaction with other salamanders. In a series of experimental trials, Muletz-Wolz et al. (2019) found infection with *Batrachochytrium dendrobatidis* increased the presence of antifungal OTUs within *Plethodon cinereus* microbiomes, suggesting skin infections can have lasting effects on an individual's cutaneous microbiome. While our comparisons of salamanders to their immediate environment is informative, future research should also include experimental manipulation of individuals to determine the effects of life history and abiotic features on the species composition of cutaneous microbial communities.

Management Implications

One of the most promising management strategies to slow potential *Bsal* outbreak is manipulation of the cutaneous microbiome through antifungal probiotics. In their list of probiotic screening recommendations, Woodhams et al. (2014) have suggested components of probiotic mixtures be chosen from locally present culturable microbiota, have the capacity to inhibit the pathogen in isolation, co-culture, and the environment, and resist any host immune defenses. We argue for additional consideration of the method in which hosts incorporate OTUs from the environment into their cutaneous microbiomes. Our results support a view of cutaneous microbiomes as functioning communities that are unlikely to incorporate new species without new resources (i.e., novel niches) or a significant disturbance. In their keystone-probiotic hypothesis, Bletz et al. (2013) introduced the concept of a rare yet impactful keystone bacterium that may affect significant changes in the community structure of an organism's microbiome when introduced. This concept is similar the shift in species composition within plant or animal communities following the introduction of invasive species (Maskell et al. 2006; Strayer 2010). In order for probiotics to be an effective deterrent of *Bsal* spread, they must be incorporated into the microbiome quickly and before infection. As such, we recommend further development of probiotics include trials of probiotic mixtures that include keystone species, and testing of probiotic mixtures using wild-caught salamanders with intact cutaneous microbiomes.

CHAPTER 5

CONCLUSION

A common thread among the results of my studies has been the identification of previously unidentified environmental factors that likely impact the biology of *P. mississippi*. Through the use of regionally replicated landscape genetics models in Chapter 2, we now know wetland connectivity plays an important role in *P. mississippi* dispersal. Knowledge that disconnected wetlands act as resistors to gene flow (as occurred in Holly Springs National Forest) and connected wetlands act as facilitators of gene flow (as occurred in Bankhead National Forest) could prove to be pivotal should the species become one of conservation concern in the future. We also saw consistent correlation between agricultural areas and resistance to gene flow among *P. mississippi* populations. The effects of hardwoods and manmade structures on salamander gene flow were mixed in Chapter 2, and both land use classes correlated with both facilitation of gene flow and resistance to gene flow under different conditions. As such, the effects of these land use types likely warrant further research. The results of Chapter 3 further emphasized the potential impact of patch configuration on the way landscape variables affect gene flow. The results also led to the identification of differences in the relationship between gene flow and agriculture in two different study areas that may be due to different agricultural uses. Within one of these study areas, they also illustrated a difference in the way short distance and long distance dispersers perceive agricultural areas, showing short distance dispersers are likely hesitant to move through agricultural areas, whereas long distance

dispersers move readily through them. These findings suggest there is no “one size fits all” approach to management of this species, and the configuration of land use types is instrumental in the movement of individuals across the landscape.

Chapter 4 led to a deeper understanding of the cutaneous microbiome of *P. mississippi*. There was no evidence of effect of kinship on salamander cutaneous microbiomes, suggesting cutaneous microbial communities are likely a product of an individuals’ environment. The species harbored a number of known antifungal OTUs within its microbiome. However, a number of antifungal OTUs were significantly more abundant in the microbiomes of individuals’ immediate environment than on their skin.

Staying true to the hallmark of scientific investigation, the research within this dissertation has led to even more questions regarding the transferability and methodology of landscape genetics analyses, as well as the relationship between amphibian cutaneous microbiomes and their environment. The results also highlight the importance of using repeated analyses, whether they are regional replicates or created via hierarchical thinning of datasets, to generate a more complete view of the relationship between environment and wildlife species. Together, the results of Chapters 2 and 3 can be combined to provide recommendations for the design of future landscape genetics studies. The field of landscape genetics was originally introduced as an intuitive method for describing patterns of genetic structure among focal species populations (Manel et al. 2003). At their simplest, landscape genetics studies seek to answer the specific question of if and how environmental variables restrict gene flow. However, studies using replication, such as those described in Chapters 2 and 3, highlight the potential predictive

power of landscape genetics inferences. By conducting repeated analyses across multiple habitat configurations or sampling densities, conservation practitioners can develop a suite of information about the way focal species respond to environmental change. This broader knowledge of the relationships between wildlife and their environment can then be incorporated into dynamic, adaptive management plans that are able to address the challenges presented by an ever-changing planet.

REFERENCES

Adamack AT, Gruber B (2014) PopGenReport: simplifying basic population genetic analyses in R. *Methods Ecol Evol* 5:384-387.

Allendorf FW, Luikart G, Aitken SN (2013) *Conservation and the genetics of populations*, 2nd ed. Wiley-Blackwell, New Jersey, USA.

Angelone S, Kienast F, Holderegger, R (2011) Where movement happens: scale-dependent landscape effects on genetic differentiation in the European tree frog. *Ecography* 34:714–722.

Bahrndorff S, Alemu T, Alemneh T, Lund Nielsen J (2016) The Microbiome of Animals: Implications for Conservation Biology. *Int J Genomics* 2016:1–7.

Bailey LL, Simons TR, Pollock KH Z (2004) Estimating detection probability parameters for plethodon salamanders using the robust capture-recapture design. *J of Wildl Manage* 68:1–13.

Bar-David S, Segev O, Peleg N, Hill N, Templeton AR, Schultz CB, Blaustein L (2007) Long-distance movements by fire salamanders (*Salamandra atra*) and implications for habitat fragmentation. *Isr J Ecol Evol* 53:143–159.

Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Stat Softw* 67:1-48.

Becker MH, Harris RN (2010) Cutaneous bacteria of the redback salamander prevent morbidity associated with a lethal disease. *PLoS One* 5:1–6.

Becker CG, Longo AV, Haddad CFB, Zamudio KR (2017) Land cover and forest connectivity alter the interactions among host, pathogen and skin microbiome. *Proc R Soc B Biol Sci* 284:

Beier P, Majka DR, Spencer WD (2008) Forks in the road: Choices in procedures for designing wildland linkages. *Conserv Biol* 22:836–851.

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Mol Ecol* 13:551–561.

Bird AK, Prado-Irwin SR, Vredenburg VT, Zink AG (2018) Skin microbiomes of California terrestrial salamanders are influenced by habitat more than host phylogeny. *Front Microbiol* 9:1–14.

Bletz MC, Loudon AH, Becker MH, Bell SC, Woodhams DC, Minbiole KPC, Harris RN (2013) Mitigating amphibian chytridiomycosis with bioaugmentation: Characteristics of effective probiotics and strategies for their selection and use. *Ecol Lett* 16:807–820.

Braunisch V, Home R, Pellet J, Arlettaz R (2012) Conservation science relevant to action: A research agenda identified and prioritized by practitioners. *Biol Conserv* 153:201–210.

Brucker RM, Harris RN, Schwantes CR, Gallaher TN, Flaherty DC, Lam BA, Minbiole KP (2008) Amphibian chemical defense: antifungal metabolites of the microsymbiont *Janthinobacterium lividum* on the salamander *Plethodon cinereus*. *J Chem Ecol* 34:1422-1429.

Bsal Task Force. (2019) North American strategic plan to control invasions of the lethal salamander pathogen Bsal. <http://www.salamanderfungus.org/>

Cabe PR, Page RB, Hanlon TJ, Aldrich ME, Connors L, Marsh DM (2007) Fine-scale population differentiation and gene flow in a terrestrial salamander (*Plethodon cinereus*) living in continuous habitat. *Hered* 98:53–60.

Castillo JA, Epps CW, Jeffrees MR, Ray C, Rodhouse TJ, Schwalm D (2016) Replicated landscape genetic and network analyses reveal wide variation in functional connectivity for American pikas. *Ecol Appl* 26:1660–1676.

Cleary KA, Waits LP, Finegan B (2017) Comparative landscape genetics of two frugivorous bats in a biological corridor undergoing agricultural intensification. *Mol Ecol* 26:4603–4617.

Collins JP, Storfer A (2003) Global amphibian declines: sorting the hypotheses. *Divers Distrib* 9:89–98.

Connette GM, Semlitsch RD (2015) A multistate mark-recapture approach to estimating survival of PIT-tagged salamanders following timber harvest. *J Appl Ecol* 52:1316–1324.

Corlett RT (2015) The Anthropocene concept in ecology and conservation. *Trends Ecol. Evol.* 30:36–41.

Costanzi J-M, Mège P, Boissinot A, Isselin-Nondedeu F, Guérin S, Lourdais O, Trochet A, Le Petitcorps Q, Legrand A, Varenne F, Grillet P, Morin-Pinaud S, Picard D (2018) Agricultural landscapes and the Loire River influence the genetic structure of the marbled newt in Western France. *Sci Rep* 8:14177.

Coulon A, Cosson JF, Angibault JM, Cargnelutti B, Galan M, Morellet N, Petit E, Aulagnier S, Hewison JM (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: An individual-based approach. *Mol Ecol* 13:2841–2850.

Cushman SA, Landguth EL (2010) Scale dependent inference in landscape genetics. *Landsc Ecol* 25:967–979.

Cox K, Maes J, Van Calster H, and Mergeay J (2017) Effect of the landscape matrix on gene flow in a coastal amphibian metapopulation. *Conserv Genet* 18:1359–1375.

Cushman SA (2006) Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biol Conserv* 128:231–240.

Cushman SA, Shirk A, and Landguth EL (2012) Separating the effects of habitat area, fragmentation and matrix resistance on genetic differentiation in complex landscapes. *Landsc Ecol* 27:369–380.

Cushman S A, Shirk A J, Landguth EL (2013) Landscape genetics and limiting factors. *Conserv Genet* 14: 263–274.

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 22: 1–20.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361.

Ellison S, Rovito S, Parra-Olea G, Vásquez-Almazán C, Flechas SV, Bi K, Vredenburg VT (2018) The Influence of Habitat and Phylogeny on the Skin Microbiome of Amphibians in Guatemala and Mexico. *Microb Ecol* 78:257-267.

Epperson BK (2010) Spatial correlations at different spatial scales are themselves highly correlated in isolation by distance processes. *Mol Ecol Resour* 10:845–853.

Epps CW, Castillo JA, Schmidt-Küntzel A, du Preez P, Stuart-Hill G, Jago M, and Naidoo R (2013a) Contrasting historical and recent gene flow among African buffalo herds in the Caprivi Strip of Namibia. *J Hered* 104:172–181.

Epps CW, Wasser SK, Keim JL, Mutayoba, BM, Brashares, JS (2013b) Quantifying past and present connectivity illuminates a rapidly changing landscape for the African elephant. *Mol Ecol* 22:1574–1588.

ESRI (2011) ArcGIS Desktop: Release 10.2. Environmental Systems Research Institute, Redlands, CA, USA.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol* 14:2611–2620.

Fahrig L, Merriam G (1985) Habitat patch connectivity and population survival. *Ecol* 66:1762–1768.

Fierer N (2017) Embracing the unknown: Disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* 15:579–590.

Galpern P, Manseau M, Wilson P (2012) Grains of connectivity: Analysis at multiple spatial scales in landscape genetics. *Mol Ecol* 21:3996–4009.

Gastón A, Blázquez-Cabrera S, Garrote G, Mateo-Sánchez PB, Simón MA, Saura S (2016) Response to agriculture by a woodland species depends on cover type and behavioural state: insights from resident and dispersing Iberian lynx. *J Appl Ecol* 53:814–824.

Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of extinction. In: *Conservation Biology: The Science of Scarcity and Diversity*. Pp. 19–34. Sinauer Associates, Sunderland, MA, USA.

Goldberg CS, Waits LP (2010) Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Mol Ecol* 19:3650–3663.

Grant EHC, Muths E, Katz RA, Canessa S, Adam MJ, Ballard JR, Berger L, Briggs CJ, Coleman J, Gray MJ, Harris MC, Harris RN, Hossack B, Huyvaert KP, Kolby JE, Lips KR, Lovich RE, McCallum HI, Mendelson III JR, Nanjappa P, Olson DH, Powers JG, Richgels KLD, Russell RE, Schmidt BR, Spitzen-van der Sluijs A, Watry MK, Woodhams DC, White CL (2016) Salamander chytrid fungus (*Batrachochytrium salamandrivorans*) in the United States—Developing research, monitoring, and management strategies: U.S. Geological Survey Open-File Report 2015–1233.

Graves TA, Beier P, Royle JA (2013) Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Mol Ecol* 22:3888–3903.

Griffiths SM, Harrison XA, Weldon C, Wood MD, Pretorius A, Hopkins K, Fox G, Preziosi RF, Antwis, RE (2018) Genetic variability and ontogeny predict microbiome structure in a disease-challenged montane amphibian. *ISME J* 12:2506–2517.

Gruber J, Brown G, Whiting MJ, Shine R (2017) Geographic divergence in dispersal-related behaviour in cane toads from range-front versus range-core populations in Australia. *Behav Ecol Sociobiol* 71:38.

Haran JM, Rossi JP, Pajares J, Bonifacio L, Naves P, Roques A, Roux G (2017) Multi-scale and multi-site resampling of a study area in spatial genetics: implications for flying insect species. *PeerJ* 5:e4135.

Hartmann SA, Steyer K, Kraus RHS, Segelbacher G, Nowak C (2013) Potential barriers to gene flow in the endangered European wildcat (*Felis silvestris*). *Conserv Genet* 14:413–426.

Hebblewhite M, Haydon DT (2010) Distinguishing technology from biology: A critical review of the use of GPS telemetry data in ecology. *Philos Trans R Soc B: Biol Sci* 365:2303–2312.

Highton, R (1989) Biochemical evolution in the slimy salamanders of the *Plethodon glutinosus* complex in the eastern United States. Part 1. Geographic protein variation. III. *Biol. Monographs*. 57:1–78.

Holderegger R, Di Giulio M (2010) The genetic effects of roads: A review of empirical evidence. *Basic Appl Ecol* 11:522–531.

Hurvich CM, Tsai C-L (1989) Regression and time series model selection in small samples. *Biometrika* 76:297–307.

Jackson CR, Stone BWG, Tyler HL (2015) Emerging perspectives on the natural microbiome of fresh produce vegetables. *Agriculture* 5:170-187.

Johnson DH (2002) The importance of replication in wildlife research. *J Wildl Manage* 66:919–932.

Kalinowski ST, Wagner AP, Taper ML (2006) ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* 6:576–579.

Keeley ATH, Beier P, Gagnon JW (2016) Estimating landscape resistance from habitat suitability: effects of data source and nonlinearities. *Landsc Ecol* 31:2151–2162.

Keeley ATH, Beier P, Keeley BW, Fagan ME (2017) Habitat suitability is a poor proxy for landscape connectivity during dispersal and mating movements. *Landsc Urban Plan* 161:90–102.

Keitt TH, Urban DL, Milne BT (1997) Detecting critical scales in fragmented landscapes. *Conserv Ecol* 1:4.

Keller D, Holderegger R, van Strien MJ (2013) Spatial scale affects landscape genetic analysis of a wetland grasshopper. *Mol Ecol* 22:2467–2482.

Keller D, Holderegger R, van Strien MJ, Bolliger J (2014) How to make landscape genetics beneficial for conservation management? *Conserv Genet* 16:503–512.

Koenig WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends Ecol Evol* 11:514–517.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120.

Landguth EL, Fedy BC, Oyler-Mccance SJ, Garey AL, Emel SL, Mumma M, Wagner HH, Fortin MJ, Cushman SA (2012) Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Mol Ecol Resour* 12:276–284.

Langwig KE, Voyles J, Wilber MQ, Frick WF, Murray KA, Bolker BM, Collins JP, Cheng TL, Fisher MC, Hoyt JR, Lindner DL, McCallum HI, Puschendorf R, Rosenblum EB, Toothman M, Willis CKR, Briggs CJ, Kilpatrick AM (2015) Context-dependent conservation responses to emerging wildlife diseases. *Front Ecol Environ* 13:195–202.

Loudon AH, Woodhams DC, Parfrey LW, Archer H, Knight R, McKenzie V, Harris RN (2014) Microbial community dynamics and effect of environmental microbial reservoirs on red-backed salamanders (*Plethodon cinereus*). ISME J 8:830-840.

Lowe WH (2010) Explaining long-distance dispersal : effects of dispersal distance on survival and growth in a stream salamander. Ecol 91:3008–3015.

Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. 71:8228-8235.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18:189–197.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220.

Maskell LC, Firbank LG, Thompson K, Bullock JM, Smart SM (2006) Interactions between non-native plant species and the floristic composition of common habitats. J Ecol 94:1052–1060.

Mayor SJ, Schneider DC, Schaefer JA, Mahoney SP (2009) Habitat selection at multiple scales. Écoscience 16:238–247.

McGarigal K, Cushman SA, and Ene E (2012) FRAGSTATS v4: spatial pattern analysis program for categorical and continuous maps. Computer software program produced by the authors at the University of Massachusetts, Amherst.

<http://www.umass.edu/landeco/research/fragstats/fragstats.html>

McMurdie PJ, Holmes S (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8:e61217.

Miller B, Soulé ME, Terborgh J (2014) ‘New conservation’ or surrender to development? Anim Conserv 17:509–515.

Moseley KR, Castleberry SB, Ford WM (2004) Coarse woody debris and pine litter manipulation effects on movement and microhabitat use of *Ambystoma talpoideum* in a *Pinus taeda* stand. For Ecol Manage 191:387–396.

Muletz CR, Myers JM, Domangue RJ, Herrick JB, Harris RN (2012) Soil bioaugmentation with amphibian cutaneous bacteria protects amphibian hosts from infection by *Batrachochytrium dendrobatidis*. Biol Conserv 152:119–126.

Muletz Wolz CR, Yarwood SA, Campbell Grant EH, Fleischer RC, Lips KR (2018) Effects of host species and environment on the skin microbiome of Plethodontid salamanders. J Anim Ecol 87:341–353.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2019) vegan: Community Ecology Package. R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>

Ousterhout BH, Semlitsch RD (2018) Effects of conditionally expressed phenotypes and environment on amphibian dispersal in nature. *Oikos* 127:1142–1151.

Ovaska K (1988) Spacing and movements of the salamander *Plethodon vehiculum*. *Herpetologica* 44:377–386.

Peakall R, Smouse PE (2012) GenAIEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research: an update. *Bioinformatics* 28:2537–2539.

Peterman WE, Connette GM, Semlitsch RD, Eggert LS (2014) Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Mol Ecol* 23:2402–2413.

Peterman WE (2018) ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods Ecol Evol* 9:1638–1647.

Petranka JW (1998) Salamanders of the US and Canada. Smithsonian Institution, Washington DC, USA.

Prado-Irwin SR, Bird AK, Zink AG, Vredenburg VT (2017) Intraspecific variation in the skin-associated microbiome of a terrestrial salamander. *Microb Ecol* 1-12.

Price MN, Dehal PS, Arkin AP (2010) FastTree 2 -- Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, 5:e9490.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genet* 155:945–959.

Prunier JG, Kaufman B, Léna JP, Fenet S, Pompanon F, Joly P (2014) A 40-year-old divided highway does not prevent gene flow in the alpine newt *Ichthyosaura alpestris*. *Conserv Genet* 15:453–468.

Prosser J, Bohannan B, Curtis T, Ellis RJ, Firestone MK, Freckleton RP, Green JL, Green LE, Killham K, Lennon JJ, Osborn AM, Solan M, van der Gast CJ, Young PW (2007) The role of ecological theory in microbial ecology. *Nat Rev Microbiol* 5:384–392.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucl Acids Res* 41:D590-D596

R Core Team. (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Heredit* 86:248–249.

Resasco J, Bruna EM, Haddad NM, Banks-Leite C, Margules CR (2017) The contribution of theory and experiments to conservation in fragmented landscapes. *Ecography (Cop)* 40:109–118.

Richardson JL, Brady SP, Wang IJ, Spear SF (2016) Navigating the pitfalls and promise of landscape genetics. *Mol Ecol* 25:849–863.

Richgels KLD, Russell RE, Adams MJ, White LC, Grant EHC (2016) Spatial variation in risk and consequence of *Batrachochytrium salamandrivorans* introduction in the USA. *R Soc Open Sci* 3:150616.

Riecken U, Raths U (1996) Use of radio telemetry for studying dispersal and habitat use of *Carabus coriaceus* L. *Annales Zoologici Fennici* 33:109–116.

Riedel BL, Russell KR, Ford WM, O’Neill KP, Godwin HW (2008) Habitat relationships of eastern red-backed salamanders (*Plethodon cinereus*) in Appalachian agroforestry and grazing systems. *Agric Ecosyst Environ* 124:229–236.

Robertson JM, Murphy MA, Pearl CA, Adams MJ, Páez-Vacas MI, Haig SM, Pilliod DS, Storfer A, Funk WC. (2018) Regional variation in drivers of connectivity for two frog species (*Rana pretiosa* and *R. luteiventris*) from the U.S. Pacific Northwest. *Mol Ecol* 27:3242–3256.

Robinson MD, McCarthy DJ, Smyth GK (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139-140

Row JR, Oyler-McCance SJ, Fike JA, O'Donnell MS, Doherty KE, Aldridge CL, Bowen ZH, Fedy BC (2015) Landscape characteristics influencing the genetic structure of greater sage-grouse within the stronghold of their range: a holistic modeling approach. *Ecol Evol* 5:1955–1969.

Row JR, Knick ST, Oyler-McCance SJ, Loughheed SC, Fedy BC (2017) Developing approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecol Evol* 7:3751–3761.

Salmerón M, Gbur EE, Bourland FM, Buehring NW, Earnest L, Fritschi FB, Golden BR, Hathcoat D, Lofton J, Thompson McClure A, Miller TD, Neely C, Shannon G, Udeigwe, TK, Verbree DA, Vories ED, Wiebold WJ, Purcell LC (2016) Yield response to planting date among soybean maturity groups for irrigated production in the US Midsouth. *Crop Sci* 56:747–759.

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF

(2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541.

Shirk AJ, Cushman SA (2014) Spatially-explicit estimation of Wright's neighborhood size in continuous populations. *Front Ecol Evol* 2:62.

Shirk AJ, Landguth EL, Cushman SA (2017) A comparison of individual-based genetic distance metrics for landscape genetics. *Mol Ecol Resour* 17:1308–1317.

Shirk AJ, Landguth EL, Cushman SA (2018) A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Mol Ecol Resour* 18:55–67.

Short Bull RA, Cushman SA, Mace R, Chilton T, Kendall KC, Landguth EL, Schwartz MK, McKelvey K, Allendorf FW, Luikart G (2011) Why replication is important in landscape genetics: American black bear in the Rocky Mountains. *Mol Ecol* 20:1092–1107.

Smith WH, and Rissler LJ (2010) Quantifying disturbance in terrestrial communities: Abundance-Biomass Comparisons of herpetofauna closely track forest succession. *Restor Ecol* 18:195–204.

Smith WH, Wooten JA, Camp CD, Stevenson DJ, Jensen JB, Turner M, Alexander NR (2018) Genetic divergence correlates with the contemporary landscape in populations of slimy

salamander (*Plethodon glutinosus* species complex) across the lower Piedmont and Coastal Plain of the southeastern United States. *Can J Zool* 96:1244–1254.

Sork VL, Waits L (2010) Contributions of landscape genetics - approaches, insights, and future potential. *Mol Ecol* 19:3489–3495.

Spatola BN, Peterman WE, Stephens NT, Connette GM, Shepard DB, Kozak KH, Selmlitsch RD, Eggert LS (2013) Development of microsatellite loci for the western slimy salamander (*Plethodon albagula*) using 454 sequencing. *Conserv Genet Resour* 5:267–270.

Spitzen-van der Sluijs A, Martel A, Asselberghs J, Bales EK, Beukema W, Bletz MC, Dalbeck L, Goverse E, Kerres A, Kinet T, Kirst K (2016) Expanding distribution of lethal amphibian fungus *Batrachochytrium salamandrivorans* in Europe. *Emerg Infect Dis* 22:1286-1288.

Steffen W, Crutzen J, McNeill JR (2007) The Anthropocene: are humans now overwhelming the great forces of Nature? *Ambio* 36:614–21.

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? *Mol Ecol* 19:3496–3514.

Strayer DL (2010) Alien species in fresh waters : ecological effects, interactions with other stressors, and prospects for the future. 55:152–174.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends Ecol Evol* 15:199–203.

Tarroso P, Carvalho SB, Velo-Anton G (in press) PHYLIN 2.0: extending the phylogeographic interpolation method to include uncertainty and user-defined distance metrics. *Mol Ecol Resour.*

Trainor AM, Walters JR, Morris WF, Sexton J, Moody A (2013) Empirical estimation of dispersal resistance surfaces: A case study with red-cockaded woodpeckers. *Landsc Ecol* 28:755–767.

U.S. Department of Agriculture Forest Service. (2004) National Forests in Alabama: Revised Land and Resource Management Plan. Region 8, Atlanta, GA. January 2004.

U.S. Department of Agriculture Forest Service. (2012) National Forests in Mississippi: Revised Land and Resource Management Plan. Region 8, Atlanta, GA. August 2014.

Van Dyck H, Baguette M (2005) Dispersal behaviour in fragmented landscapes: routine or special movements? *Basic Appl Ecol* 6:535–545.

van Etten J (2017) R package gdistance: distances and routes on geographical grids. *J Stat Softw* 76(13).

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538.

Vergara M, Cushman SA, Ruiz-González A (2017) Ecological differences and limiting factors in different regional contexts: landscape genetics of the stone marten in the Iberian Peninsula. *Landsc Ecol* 32:1269–1283.

Wagner HH, Holderegger R, Werth S, Gugerli F, Hoebee SE, Scheidegger C (2005) Variogram analysis of the spatial genetic structure of continuous populations using multilocus microsatellite data. *Genetics* 169:1739–1752.

Walke JB, Becker MH, Loftus SC, House LL, Cormier G, Jensen RV, Belden LK (2014) Amphibian skin may select for rare environmental microbes. *ISME J* 8:2207–2217.

Wang IJ (2009) Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). *Mol Ecol* 18:3847–3856.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol.* 73:5261-5267.

Wells KD, Wells RA (1976). Patterns of movement in a population of the slimy salamander, *Plethodon glutinosus*, with observations on aggregations. *Herpetologica* 156–162.

Wenger SJ, Olden JD (2012) Assessing transferability of ecological models: An underappreciated aspect of statistical validation. *Methods Ecol Evol* 3:260–267.

Woodhams DC, Brandt H, Baumgartner S, Kielgast J, Küpfer E, Tobler U, Davis LR, Schmidt BR, Bel C, Hodel S, Knight R, McKenzie V (2014) Interacting symbionts and immunity in the amphibian skin mucosome predict disease risk and probiotic effectiveness. *PLoS One* 9:4 e96375.

Woodhams DC, Alford RA, Antwis RE, Archer H, Becker MH, Belden LK, Bell SC, Bletz M, Daskin JH, Davis LR, Flechas SV, Lauer A, Gonzalez A, Harris RN, Holden WM, Hughey MC, Ibáñez R, Knight R, Kueneman J, Rabemananjara F, Reinert LK, Rollins-Smith LA, Roman-Rodriguez F, Shaw SD, Walke JB, McKenzie V (2015) Antifungal isolates database of amphibian skin-associated bacteria and function against emerging fungal pathogens. *Ecology* 96:595–595.

Yuan ML, Dean SH, Longo AV, Rothermel BB, Tuberville TD, Zamudio KR (2015) Kinship, inbreeding and fine-scale spatial structure influence gut microbiota in a hindgut-fermenting tortoise. *Mol Ecol* 24:2521–2536.

Zeller KA, McGarigal K, Whiteley AR (2012) Estimating landscape resistance to movement: A review. *Landsc Ecol* 27:777–797.

APPENDIX

Summary of Laboratory Methods

We assessed 27 loci developed by Spatola et al. (2013), and found nine loci that reliably amplified for *Plethodon mississippi* from Holly Spring National Forest, Mississippi (HSNF) and Bankhead National Forest, Alabama (BNF). Of these nine loci, PLAL_B8DRY was monomorphic in HSNF, and PLAL_402 failed to amplify in over half the BNF samples.

Accordingly, each these two loci were used only for screening samples from the forest region in which the locus was both polymorphic and amplified reliably. Polymerase chain reactions (PCRs) were performed in 15 μ L volumes, each containing the reagents listed in Table A1. Several samples failed to amplify and in these cases, amplifications were performed in 15 μ L volumes, each containing the reagents listed in Table A2.

PCR products were sent to Yale University's DNA Analysis Facility on Science Hill for fragment analysis. A ROX-500 size standard was used to estimate allele lengths. We used the software Geneious v.9.1.2 (<http://www.geneious.com>) to examine the resulting .fsa files and score genotypes. Allele sizes were assigned using bins whose bounds were set -2 base pairs (bp) and +1 bp the size of each allele (all loci contained tetranucleotide repeat motifs). Each .fsa file was viewed and scored individually. MICRO-CHECKER v.2.2.3 (Van Oosterhout et al. 2004) was used to identify alleles sizes that were indicative of stutter peaks or errors created when recording allele sizes (Table A3).

Table A1: Reagent list for standard PCRs.

Reagent	Concentration	Volume
5X <i>GoTaq</i> Flexi Buffer (Promega)	-	3.0 μ L
MgCl ₂ (Promega)	25mM	1.2 μ L
dNTPs (Promega)	1.25mM	2.4 μ L
Bovine Serum Albumin (New England Biolabs)	10mg/mL	0.75 μ L
dH ₂ O	-	3.25 μ L
Forward primer (with 5' M13 tail)	1 μ M	0.75 μ L
Reverse primer	10 μ M	0.75 μ L
Universal M13 primer (with 5' HEX fluorescent label)	10 μ M	0.75 μ L
<i>GoTaq</i> (Promega)	5U/ μ L	0.15 μ L
Genomic DNA	Approx. 40ng/ μ L	2.0 μ L

Table A2. Reagent list for PCRs of difficult samples using Qiagen's Type-It Microsatellite PCR Kit.

Reagent	Concentration	Volume
Type-It Microsatellite PCR Kit Master Mix (Qiagen)	-	7.5 μ L
dH ₂ O	-	3.75 μ L
Forward primer (with 5' M13 tail)	1 μ M	0.75 μ L
Reverse primer	10 μ M	0.75 μ L
Universal M13 primer (with 5' HEX fluorescent label)	10 μ M	0.75 μ L
Genomic DNA	Approx. 40ng/ μ L	1.5 μ L

Table A3. Microsatellite primers successfully amplified from Spatola et al. (2013).

Locus names were maintained from the original literature. Annealing temperature in °C (T_a), number of alleles N_A , and allelic size range in basepairs by chapter are also specified. The sequence of a 5' tail used to attach the fluorescently labeled M13 primer is in bold at the 5' end of each forward primer.

Locus name	Repeat motif	Primer sequence (5' to 3')	T_a (°C)	Chapter II N=220		Chapter III N=183		Chapter IV N=45	
				N_A	Allele size range (bp)	N_A	Allele size range (bp)	N_A	Allele size range (bp)
PG_43M	AATG	F: TCCCAGT CACGACCGTAGTCATTGTCAAGCTTGCGC R: GGGAGCTTGGATCAGGAAAG	55	25	106-202	9	114-150	4	226-262
PG_POG	AATG	F: TCCCAGT CACGACCGTACCTGTATTTCACGCTGCAC R: CTGCACCTCTCACCCCTACTG	60	22	158-270	9	238-266	7	112-142
PG_QWZ	AATG	F: TCCCAGT CACGACCGTTCGTCTGATTATTGGCGTGC R: ACCTATCTCATCCCACTGC	60	12	143-195	5	163-179	3	167-179
PG_V58	AATG	F: TCCCAGT CACGACCGTGTGCCACCTTGTTTCCTG R: TTGTGAGTCTCCTGCCCTTG	60	13	146-202	9	146-186	4	146-170
PLAL_241	AATG	F: TCCCAGT CACGACCGTCAAAGGTAGGCAATGGTCTCG R: ACGAGCTAGACCTCTATTITGGG	55	14	203-315	8	203-255	4	215-235
PLAL_402	ATCC	F: TCCCAGT CACGACCGTAGTGGTAGGGGAGATGGATG R: TGGACTGTGCTTCTTGTGC	60	32	111-267	32	111-267	19	111-227
PLAL_545b	AGAT	F: TCCCAGT CACGACCGTGGGCCTGGAGCATTACATA R: GCTTAGTGCAAGGTGCTTCC	60	36	178-394	28	178-394	19	302-394
PLAL_615	AGAT	F: TCCCAGT CACGACCGTCTAAAGAGCACGGGACAGAG R: TATGAGTCCGATCGGTGAGC	60	35	150-314	27	154-314	18	190-262
PLAL_B8DRY	AGAT	F: TCCCAGT CACGACCGTGTGGACAGTGGGATCAAGG R: TGATGGAGCCAGAGACATG	60	29	182-346	-	-	-	-

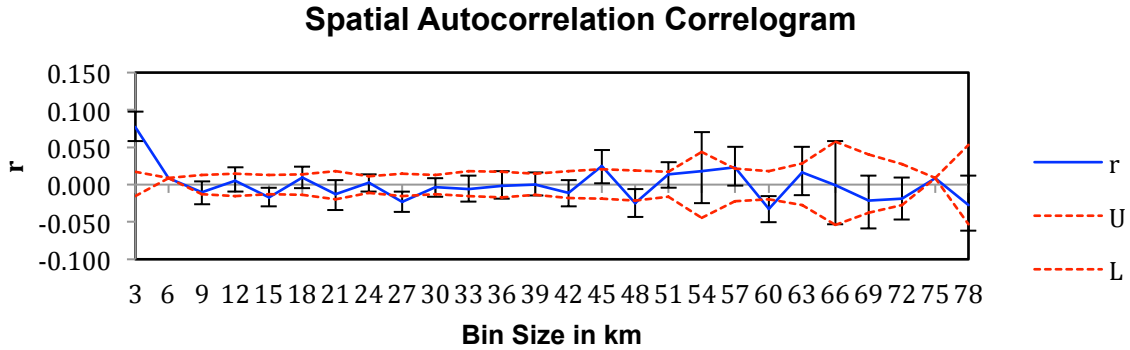


Figure A1. Spatial autocorrelation analysis for *P. mississippi* individuals in Holly Springs National Forest (HSNF) generated using GenAIEx 6.503.

The x-axis displays correlogram bins of 3km, and the y-axis displays the spatial autocorrelation coefficient “r”. The red dotted lines indicated the upper (U) and lower (L) confidence intervals and 95% and 5% respectively. A heterogeneity test for correlogram significance was significant ($p < 0.001$). The x-intercept, or genetic neighborhood size, was 7.4km.

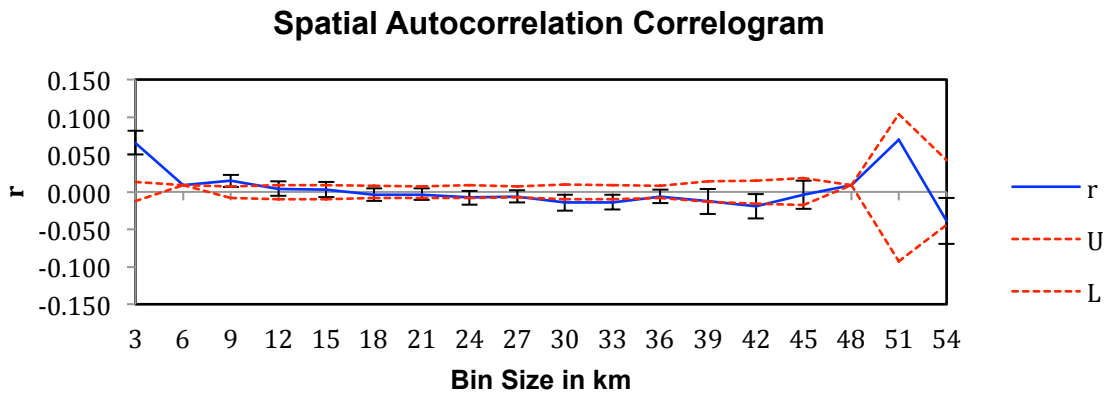


Figure A2. Spatial autocorrelation analysis for *P. mississippi* individuals in Bankhead National Forest (BNF) generated using GenAIEx 6.503.

The x-axis displays correlogram bins of 3km, and the y-axis displays the spatial autocorrelation coefficient “r”. The red dotted lines indicated the upper (U) and lower (L) confidence intervals and 95% and 5% respectively. A heterogeneity test for correlogram significance was significant ($p < 0.001$). The x-intercept, or genetic neighborhood size, was 16.3km.

Table A4: Percent of each land use class found in the large and small study area within Holly Springs National Forest.

Land use class	Large study area (630 km ²)	Small study area (256 km ²)
Agriculture	8.2%	9.9%
Hardwoods	31.5%	25.2%
Pine	41.8%	45.9%
Manmade structures	4.6%	4.6%
Water	10.6%	11.4%
Wetlands	3.3%	3.0%

Table A5. OTUs that comprise the salamander core microbiome.

The listed OTUs are present in at least 80% of salamander microbiome samples. Bolded OTUs have >97% identity with a bacterial isolate listed as “inhibitory” in the Amphibian Skin Isolates database by Woodhams et al. (2015). The abbreviation “uncl” represents the term unclassified.

OTU#	Phylum	Class	Order	Family	Genus
00002	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium
00003	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
00004	Proteobacteria	Betaproteobacteria	Burkholderiales	Burkholderiaceae	Burkholderia
00005	Proteobacteria	Alphaproteobacteria	Rhizobiales	Roseiarcaceae	Roseiarcus
00007	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales_uncl.	Rhizobiales_uncl.
00012	Planctomycetes	Planctomycetia	Planctomycetales	Planctomycetaceae	Planctomycetaceae_uncl.
00024	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales_uncl.	Rhizobiales_uncl.
00046	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methyllobacteriaceae	Methyllobacterium
00050	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Enterobacteriaceae_uncl.
00062	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas
00122	Proteobacteria	Betaproteobacteria	Burkholderiales	Comamonadaceae	Comamonadaceae_uncl.
00193	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia

Table A6. Classifications of OTUs associated with cluster B.

The listed OTUs have a score of >0.2 along axis 1 of a Principal Coordinates Analysis using UNIFRAC distance, placing them in the same oriented space as several salamander microbiome samples that have clustered apart from soil microbiome samples. All of the samples below are present in at least 25% of microbiome samples. The abbreviation “uncl.” represents the term unclassified. The bolded OTUs have >97% identity with a bacterial isolate listed as “inhibitory” in the Amphibian Skin Isolates database by Woodhams et al. (2015).

OTU #	Axis 1	Phylum	Class	Order	Family	Genus
00074	0.263	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.
00037	0.259	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.
00054	0.258	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales_uncl.	Rhizobiales_uncl.
00026	0.258	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Hymenobacter
00034	0.257	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.
00045	0.255	Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	Massilia
00030	0.255	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.	Bacteria_uncl.
00013	0.255	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Hymenobacter
00025	0.255	Bacteroidetes	Cytophagia	Cytophagales	Cytophagaceae	Hymenobacte
00067	0.255	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiales_uncl.	Rhizobiales_uncl.
00046	0.252	Proteobacteria	Alphaproteobacteria	Rhizobiales	Methylobacteriaceae	Methylobacterium
00040	0.251	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	Microbacteriaceae_uncl.
00089	0.245	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas
00123	0.243	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas

VITA

STEPHANIE BURGESS, MS, PHD

EDUCATION

University of Mississippi, Oxford, MS

Ph.D in Biology
Dissertation: *Application and Refinement of Molecular Ecology Techniques for Amphibian Conservation* 08/19

Virginia Commonwealth University, Richmond, VA

Msc. in Biology
Thesis: *Dispersal and population structure of Ambystoma opacum* 05/08
Graduate Certificate, Geographic Information Systems (GIS) 05/08
Bsc. in Biology 05/06

APPOINTMENTS

University of Mississippi, Oxford, MS

Graduate Dissertation Fellow 01/2019 – 05/2019
Graduate Teaching Assistant (Introductory Biology for majors) 08/2014 - 12/2018

Big Thicket National Preserve, Village Mills, TX

Biologist, Oil and Gas Program Manager 11/2008 - 05/2014

Texas Commission on Environmental Quality, Beaumont, TX

Environmental Investigator 07/2008 - 10/2008

Virginia Commonwealth University, Richmond, VA

Graduate Teaching Assistant 08/2006 - 05/2008

Back Bay National Wildlife Refuge, Virginia Beach, VA

Wildlife Biology Intern 05/2007 – 08/2007

RESEARCH

Burgess S, and Garrick R. Regional Replication of Landscape Genetics Analyses of the Mississippi Slimy Salamander, *Plethodon mississippi*. Landscape Ecology (In Review)

Garrick R, Banusiewicz J, **Burgess S**, Hyseni C, and Symula R. 2019. Extending phylogeography to account for lineage fusion. Journal of Biogeography 46(2): 268-278.

Garrick R, Reppel D, Morgan J, **Burgess S**, Hyseni C, Worthington R, and Ulyshen M. 2019. Trophic interactions among dead-wood-dependent forest arthropods in the southern Appalachian Mountains, USA. Food Webs 18:e00112.

Buchholz R, Banusiewicz JD, **Burgess S**, Crocker-Buta S, Eveland L, and Fuller L. 2019. Behavioural research priorities for the study of animal response to climate change. Animal Behaviour (In press)

PRESENTATIONS

Evolution 2016 Poster: *Landscape Genetics and the impact of Neighboring Landscapes: the Proximity Effect*

Ecological Society of America 2017 Conference Presentation:
Landscape Genetic Analysis of Plethodon mississippi in Holly Springs National Forest, MS

8th Annual University of Mississippi Graduate Research Symposium Presentation:
An Assessment of the Transferability of Amphibian Landscape Genetics Models
1st Place in STEM group, \$650 award

US-International Association of Landscape Ecologists 2018 Conference Presentation:
Determining the Transferability of Amphibian Landscape Genetics Models

US-International Association of Landscape Ecologists 2019 Conference Presentation:
The Effect of Sampling Scale on Landscape Genetics Analyses for Plethodon mississippi

GRANTS AND AWARDS

NASA-MSU Professional Enhancement Award 09/16 – 05/17
Funded attendance to the \$700
2018 US-International Association of Landscape Ecologists annual conference

Graduate Student Council Research Grants Program, University of Mississippi 09/16 - 05/17
An Assessment of the Transferability of Amphibian Landscape Genetics Models \$1000

Birmingham Audubon Society Walter F. Coxe Research Grant 02/17 - 02/19
Assessing the Effect of Adjacent Land Use on \$1965
Animal Movement using Landscape Genetics

PROFESSIONAL SERVICE & OUTREACH

Reviewer: *Bulletin of Entomological Research*

Oxford Elementary School Science Fair Judge (2016 and 2017)

Natchez Trace National Parkway Bioblitz Volunteer (2017)

University of Mississippi Biology Graduate Student Society Vice President (2016 and 2017)