Landscape Genetics of Ambystoma opacum in Mammoth Cave National Park

Haley Austin
Western Kentucky University, haley.austin671@topper.wku.edu

Follow this and additional works at: https://digitalcommons.wku.edu/stu_hon_theses

Part of the Biology Commons, Ecology and Evolutionary Biology Commons, and the Genetics Commons

Recommended Citation
https://digitalcommons.wku.edu/stu_hon_theses/737

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Honors College Capstone Experience/Thesis Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
LANDSCAPE GENETICS OF AMBYSTOMA OPACUM IN MAMMOTH CAVE NATIONAL PARK

A Thesis Project Presented in Partial Fulfillment of the Requirements for the Degree Bachelor of Science with Honors College Graduate Distinction at Western Kentucky University

By

Haley D. Austin

Western Kentucky University

2018

CE/T Committee:
Professor Jarrett Johnson, Advisor
Professor Melanie Autin
Professor Thomas Richmond

Approved by

Advisor
Department of Biology
ABSTRACT

Landscape genetics describes relationships between landscape variables and genetic variation in plant and animal populations. This has contributed to a better understanding of how environmental changes can affect the genetic composition and survival of a population. Over recent decades, global amphibian populations have been declining. An understanding of habitat structure and connectivity is important to consider when developing effective conservation strategies. The purpose of this study is to investigate the effect of landscape characteristics on gene flow and population structure of the marbled salamander (*Ambystoma opacum*) in Mammoth Cave National Park (MCNP). This was accomplished using ResistanceGA, an R package, to optimize the landscape and assign resistance values to five habitat types: wet deciduous forest, dry deciduous forest, coniferous forest, human influence, and water. The program used coordinate locations from 50 sample sites, pairwise genetic distances between those ponds, and GIS landscape data from the park. Preliminary results indicate that distance is the best predictor of pairwise genetic distances. Adding vegetation to the model did not significantly improve the model. Within the vegetation model, human influence is least resistant to movement, followed by water, wet deciduous forest, coniferous forest, and dry deciduous forest.

Keywords: *Ambystoma opacum*, isolation-by-distance, isolation-by-resistance, landscape genetics, Mammoth Cave, marbled salamander
ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Jarrett Johnson, for his knowledge, guidance, and patience throughout this process. I would also like to thank my second reader, Dr. Melanie Autin, for all of her help as I struggled to find a topic, and for her assistance and feedback in the rest of the project.

Thank you to the WKU Bioinformatics and Information Science Center and Samantha Thomas for technical assistance, to Kevin Martin for use of the genetic data, and to the WKU Biotechnology and Biodiversity Centers for support. The collection of genetic data was supported by a KAS Marcia Athey Grant.

I would like to thank my friends for encouraging me when I felt inadequate and for always reminding me that my worth is not rooted in my academic success. Thanks to Elizabeth, Lydia, and Lindsey, for navigating this process with me from the beginning and for inspiring me with the passion you each have for your respective fields. A special thanks to my roommate, Elizabeth, for occasionally forcing me out of bed in the morning.

To my dogs, Belle, Loofa, and Winnie, thank you for always brightening my day and providing much-needed emotional support as I navigated college and this thesis.

Finally, I would like to thank my parents for their constant love, support, and encouragement. Thank you for prioritizing my education and always encouraging me to pursue my interests. I am forever grateful to you for the example you have set for me, and for always pushing me to grow in all areas of my life.
VITA

EDUCATION

Western Kentucky University, Bowling Green, KY
Bachelor’s of Science – Mahurin Honors College Graduate

Madison Central High School, Richmond, KY

AWARDS & HONORS

Larry N. Gleason Award for Excellence in Research, WKU, 2018
Session winner, WKU Student Research Conference, 2018
Diversity Abroad Grant, WKU, 2016
Honors Travel Abroad Grant, 2016
World Topper Scholarship, 2016
1906 Founders Scholarship, WKU, 2014-2018

PROFESSIONAL MEMBERSHIPS

TriBeta National Biological Honor Society

INTERNATIONAL EXPERIENCE

Harlaxton College, Grantham, United Kingdom

January 2016-May 2016

PRESENTATIONS

TABLE OF CONTENTS

Abstract..................................................................................................................ii
Acknowledgements...............................................................................................iii
Vita..........................................................................................................................iv
List of Figures.........................................................................................................vi
List of Tables..........................................................................................................vii
Introduction...........................................................................................................1
Methods..................................................................................................................7
Results....................................................................................................................12
Discussion..............................................................................................................13
References..............................................................................................................17
LIST OF FIGURES

Figure 1. Photos of study site..........................................................................................21
Figure 2. Land cover map..................................................................................................22
Figure 3. Map of resistance surface..................................................................................23
Figure 4. Circuitscape heat map........................................................................................24
LIST OF TABLES

Table 1. Descriptions of habitat types ........................................................................25
Table 2. ResistanceGA models tested ........................................................................26
Table 3. Model comparisons .......................................................................................27
Table 4. Resistance values calculated .........................................................................28
INTRODUCTION

Amphibian populations have been declining globally over recent decades, due to threats such as habitat loss and exploitation, climate change, and disease (Houlahan et al. 2000; Stuart et al. 2004; Blaustein et al. 2011; Campbell-Grant et al. 2016). Loss of connectivity due to habitat fragmentation, whether natural or anthropogenic, can have negative impacts on gene flow, which can reduce genetic diversity (Foley 2005; Hamer and McDonnel 2008; Morten and O’Brien 2010; Sunny et al. 2014). Understanding how landscape variables affect connectivity and genetic variation in populations, allows for the development of effective conservation models (Bennett 2003; Morten and O’Brien 2010).

Loss of connectivity can lead to isolation of subpopulations, resulting in reduced genetic diversity, and consequently, the ability to adapt (Lacy 1987; Bennett 2003; Sunny et al. 2014). Smaller, more isolated populations experience reduced gene flow and are more susceptible to inbreeding and genetic drift (Lacy 1987; Sjögren 1991). With fewer genes entering the gene pool, breeding within the population can perpetuate suboptimal alleles (Allentoft 2010). Inbreeding increases the probability of an individual inheriting two copies of such deleterious alleles, and, along with genetic drift, can cause loss of heterozygosity (Lacy 1987). While lethal deleterious alleles may be easily removed from the population through natural selection, sub-lethal alleles can persist in the population,
affecting overall fitness of the individuals and potentially leading to extinction (Sjögren 1991; Allentoft 2010). Thus, understanding and preserving connectivity between subpopulations is important for local conservation efforts. There is not a universal solution to the problem of amphibian declines because the causes of decline vary across continental, regional, and local scales. Thus, the development of local conservation strategies is essential to counteracting population declines (Campbell Grant et al. 2016).

A common method for preserving connectivity within landscapes is through habitat corridors. Boundaries between undeveloped and developed habitat are often abrupt and are often referred to as ‘hard edges’ (Fenske-Crawford and Niemi 1997; Bennett 2003). Hard edges may pose physical and psychological barriers to movement (Mader 1984; Gibbs 1998). In contrast, natural boundaries are more gradual transitions, making them less resistant to movement (Fenske-Crawford and Niemi 1997; Bennett 2003). Habitat corridors are a way to facilitate movement and increase gene flow between suitable habitats through developed areas by eliminating harsh edges and providing a section of preferable habitat through the resistant area (Haddad and Tewksbury 2005). Stepping stones, which are like habitat corridors but segmented, create habitats where animals are only forced to move short distances through disturbed habitat (Saura et al. 2014; Sunny et al. 2014). Selection of habitat corridors or stepping stones depends on the level of modification of the habitat, as well as the life-history traits of the species involved (Saura et al. 2014). Preservation of connectivity in these ways maintains landscape-scale genetic diversity, facilitating adaptation and population persistence.

Landscape genetics, an interdisciplinary field that combines population genetics, landscape ecology, and spatial statistics, evaluates relationships between patterns of
genetic variation and landscape variables (Manel et al. 2003; Storfer et al. 2007). The primary purpose of landscape genetics is to develop a better understanding of how modifications to landscape, caused by factors such as changes in land use and climate change, impact genetic diversity (Manel and Holderegger 2013). It also considers whether a species will be able to adapt as such changes continue (Manel and Holderegger 2013). Landscape genetics involves the correlation of genetic discontinuities and landscape features, which is accomplished by the selection of a model to explain a genetic response (Manel et al. 2003). Potential models include isolation-by-distance and isolation-by-resistance models. Isolation-by-distance models explain genetic distances between populations with physical distance between the populations, assume homogenous, unbounded populations, and do not take into account variations in population densities or migration rates across the range (McRae 2006). Isolation-by-resistance models consider this variability and explain genetic distances between populations with landscape structure (McRae 2006). An isolation-by-slope model, for example, is an isolation-by-resistance model in which steeper slope is more resistant to gene flow (Mims et al. 2016). Isolation-by-barrier models use potential complete obstructions to gene flow, such as a fenced highway, to explain gene flow in a population (Storfer et al. 2007).

In early landscape genetics studies, the most common statistical technique used in model selection was the partial Mantel test. A classical Mantel test first identifies the correlation between pairwise genetic and geographic distance matrices (Manel and Holderegger 2013). The correlation coefficient is then compared to those derived from random permutations of one of the matrices, while the other is held constant (Raufaste
and Rousset 2001; Legendre and Fortin 2010). Under the null hypothesis that geographic distance has no effect on genetic distance, the probability of each permutation is equal. However, in a partial Mantel test, a third matrix is added with another environmental explanatory variable (Balkenhol et al. 2009). Under the null hypothesis that this third parameter has no effect on the response, each permutation would depend on the effect of geographic distance, invalidating the permutation procedure and making it impossible to know the true degrees of freedom (Raufaste and Rousset 2001; Manel and Holderegger 2013). Without knowing the degrees of freedom, measures used in model selection methods, such as $R^2$ and Akaike information criterion (AIC) cannot be used (Manel and Holderegger 2013). The Mantel test also generates high type-1 error rates, meaning it often identifies landscape variables as significant when they have no true influence on gene flow (Balkenhol et al. 2009).

The use of mixed-effects models is a potential alternative that accounts for the non-independence of the predictor variables. ResistanceGA (Peterman et al. 2014), a package for R (R Core Team 2013), fits linear mixed-effects models to the data using the maximum likelihood population effects parametrization (MLPE) to account for the non-independence (Clark et al. 2002; Peterman 2014). When compared with six other model selection methods in population genetic simulations, linear mixed-effects models with MLPE were most likely to select the correct model (Shirk et al. 2017). ResistanceGA can be used to optimize both continuous and categorical surfaces, and it has the capacity to optimize multiple surfaces simultaneously. ResistanceGA optimization is an iterative process that begins with the generation of a random population of a specified size and the calculation of least-cost paths across the landscape (Peterman 2014). Linear mixed-
effects models are fit to each member of the population, where pairwise effective distance is the independent variable, and pairwise genetic distance is the response. Models with the best objective function, either log-likelihood, AIC, or $R^2$, are selected by a genetic algorithm, and the next group of individuals is generated based on properties of the selected individuals. Iterations continue until the objective function fails to improve over a specified number of generations (Peterman 2014). While effective, genetic algorithms are randomized procedures, and should be run multiple times to confirm model selection (Scrucca 2013). The optimization process is computationally demanding, and the time required for each generation increases with the number of sample locations and size of the landscape.

**Predictions**

My objective was to investigate the effect of landscape characteristics on gene flow and population structure of the marbled salamander (*Ambystoma opacum*) in Mammoth Cave National Park. While *A. opacum* is not currently threatened, understanding the relationship between gene flow and the park’s landscape will inform the development of strategies to maintain high genetic diversity and to prevent population declines. While amphibian conservation strategies often need to be species- and landscape-specific, the results of this analysis could contribute to the development of generalized strategies where individualized approaches are not feasible.

Based on previous analyses of the data, I expect an isolation-by-distance model to best explain genetic distance between ponds. In an isolation-by-resistance model, I expect
flowing water and human development to constitute barriers with high resistance to movement and wet forest to be most conducive to movement.
METHODS

Study Species

*Ambystoma opacum* is found in a variety of habitats throughout much of the Eastern United States, including the state of Kentucky, and most commonly in wooded areas (Barbour 1971; Beane 2010). Adults are terrestrial and spend the majority of their time under rocks and logs (Barbour 1971; Beane 2010). Adult lengths range from about 3.5 to 5 inches (Barbour 1971; Beane 2010). The breeding season is from September to November, following a return from summer habitats to breeding sites, which are often natal ponds (Barbour 1971; Scott 1994). Females lay eggs on land and typically stay with the eggs until the area floods. This typically occurs in the fall, and then the eggs hatch (Beane 2010). The larvae then metamorphose the following summer (Barbour 1971).

Study Landscape

The 50 sample ponds (Figure 1) were selected from Mammoth Cave National Park (37.184461° -86.098934°). In addition to *A. opacum*, Mammoth Cave is home to 15 species of frog, and 17 species of salamander. The park lies within the south-central Kentucky karst region, and is bisected by the Green River. North of the river, exposed limestone has caused uneven topography. South of the river, insoluble rocks cover the limestone, preventing erosion. Mammoth Cave consists of forest, savannah, and small patches of prairie ecosystem and is high in plant diversity.
**Sample Collection**

As described in Martin (2013), ponds were identified using topographic map data, GIS wetlands layers, Google Earth imagery, park ranger knowledge, and random encounter. Fifty-two of the 60 ponds visited between January and April of 2012 contained *A. opacum* larvae. Two ponds were ultimately excluded from the dataset, one due to insufficient sample size and the other due to the pond’s drying up. At each pond, 12-30 larvae were collected with dip nets, and a small tail clipping of approximately 1 cm was clipped from each individual. Tail clipping is an efficient method for collection of genetic data that has been shown to have minimal effect on survival of the individual (Wilbur and Semlitsch 1990). Larvae were released immediately following clipping, and tissues were stored in 95% ethanol. Collection was supported by Kentucky Department of Fish and Wildlife permit #SC1211057 and Mammoth Cave Scientific Research and Collecting permit #MACA-2012-SCI-0001.

**Collection of Genetic Data**

As described in Martin (2013), DNA was extracted from the collected tissues using standard phenol-chloroform procedures, DNEasy Blood and Tissue Kits® (Qiagen Inc.) or ethanol precipitation methods. Genomic DNA was then screened for amplification and polymorphism of microsatellite markers at 10 *A. opacum* loci (Nunziata *et al.* 2011). The polymerase chain reaction (PCR) products from each marker were multiplexed for all individuals and were then genotyped using a 3130 Genetic Analyzer and GeneMapper® v3.7 software (Applied Biosystems, Inc.) at the Western Kentucky University Biotechnology Center, or with a 3730xl 96-capillary DNA Analyzer at the University of Georgia Genomics Facility.
Collection of Geographic Data

As described in Martin (2013), pairwise geographic distances were calculated using GPS coordinates of the pond sample sites and ArcMap (Esri, Inc.). A categorical raster surface was created with data from the United States Geological Survey (USGS 2011). Each habitat type in the raster surface was assigned a category based on type of vegetation and moisture of the habitat (Table 1). The raster layer included five landscape types: wet deciduous forest, dry deciduous forest, coniferous forest, human influence, and water (Figure 2).

Analysis with ResistanceGA

ResistanceGA was used to optimize the study landscape and to select an isolation-by-distance or an isolation-by-resistance model to describe movement between breeding ponds. ResistanceGA uses landscape data, pairwise genetic distances between sample locations, and the coordinates of the sample locations to assign resistance values to landscape variables (Peterman 2014). In this analysis, the landscape data input was a categorical raster layer created using data from the US Geological Survey from Mammoth Cave National Park. The pairwise genetic distances between ponds were input as normalized \( F_{st} \) values calculated by Martin (2013), and the coordinate locations of the 50 ponds were input as sample locations.

The linear mixed-effect models fit to the data include the pairwise genetic distance as the response variable and use fixed and random effect terms as explanatory variables. The pairwise pond ID was the random effect term considered, and fixed effect terms included distance between ponds and vegetation resistance between ponds. Models considered in the analysis included a null model, where pairwise pond ID was the only
explanatory variable; an isolation-by-distance model, which included distance in addition to pairwise pond ID; and an isolation-by-resistance model, which added vegetation as an explanatory variable (Table 2). The code following code was used to run ResistanceGA in R.

```r
library(ResistanceGA)
write.dir<-"/home/haustin/Data/
veg<-raster("/home/haustin/Data/veg.asc")
ponds<-read.csv("/home/haustin/Data/MACA_pondlocs.csv")
ponds.spatial<-SpatialPoints(ponds)
GA.inputs<-GA.prep(ASCII.dir=write.dir, method="LL", max.cat=10,
parallel=4, seed=10, quiet=T)
genetic<-read.csv("/home/haustin/Data/fstnormalizedmatrix.csv", header=FALSE)
gdist.inputs<-gdist.prep(50, samples=ponds.spatial,
response=lower(as.matrix(genetic)))
SS_RESULTS.gdist<-SS_optim(gdist.inputs=gdist.inputs,
GA.inputs=GA.inputs)
```

Each generation in ResistanceGA consisted of 30 iterations for the distance optimization and 90 for the vegetation optimization. For each iteration, a linear mixed-effects model with MLPE was created, and resistance values between 1 and 10 were assigned to each vegetation type. In Martin (2013), resistance values were assigned by hand and altered one-at-a-time until the partial Mantel $r$ value associated with the model was maximized. ResistanceGA operated similarly, but in an automated format in which one resistance value was held constant and the others were manipulated until AICc was optimized along with the distance parameter. The random number seed for the genetic algorithm was set to 10, and the algorithm was run in parallel with 4 cores using the Western Kentucky University Bioinformatics Center’s computer cluster. The computer cluster uses multiple nodes to improve the efficiency of processing. The genetic algorithm selected the models with the best log-likelihood values and generated new models in an attempt to further maximize these values. The genetic algorithm continued
to run until 25 generations passed without improvement in AICc. When sample size is small, AIC is biased toward models with more parameters. AICc corrects for this bias.

The ASCII file of habitat resistances generated by ResistanceGA was used to create a heat map using the program Circuitscape (McRae et al. 2008) to visualize movement of *A. opacum* at MCNP.
RESULTS

After 40 iterations, ResistanceGA converged on an isolation-by-distance model (Table 3). The addition of vegetation to the model did not significantly improve its ability to predict pairwise genetic distances. Log-likelihood was increased from 2995.4689 to 2998.6543, and AIC was reduced from -5982.9377 to -5989.3086. However, AICc, which corrects for bias towards models with additional parameters, was not reduced with the addition of the vegetation parameters.

In the isolation-by-resistance model, human influence was assigned the lowest resistance value (1), followed by water (1.6207), wet deciduous forest (1.6207), coniferous forest (5.8722), and dry deciduous forest (6.5647) (Table 4, Figure 3). The Circuitscape heat map generated from the selected resistances shows higher movement between ponds that are closer in proximity and separated by habitat types with lower resistances, particularly roads (Figure 4).

Analysis described in Martin (2013) also selected an isolation-by-distance model. In the resistance model, wet deciduous forest was assigned the lowest resistance value (8), followed by dry deciduous forest and human influence (11), coniferous forest (13), and water (23). While the resistance values calculated by Martin (2013) are not directly comparable to those calculated in this study due to the differing magnitude, comparisons between the relative resistance values of the habitat types were of interest.
DISCUSSION

Based on the results, isolation-by-distance model better describes gene flow of *A. opacum* at MCNP than the isolation-by-resistance model. The addition of vegetation does not significantly improve the model. Therefore, vegetation in MCNP does not significantly affect gene flow in the marbled salamander populations. This conclusion is consistent with previous analysis of these data (Martin 2013). However, the relative resistance values assigned to the habitat categories differ.

Previously, wet deciduous forest was identified as the least resistant category, followed by both dry deciduous forest and human influence (equal resistance values were assigned to these two categories), coniferous forest, and water (Martin 2013). ResistanceGA assigned resistances to forest types in the same order as before, which is consistent with habitat moisture requirements in amphibians. However, ResistanceGA determined both human influence and water to be less resistant to movement than any of the forest types.

ResistanceGA stopped running prior to reaching the expected number of iterations. It is possible that an incorrect model was selected. The results obtained from the ResistanceGA run are considered preliminary due to the uncertainty caused by the short run. However, it is possible that human influence truly is least resistant to *A. opacum* movement at Mammoth Cave. Areas affected by anthropogenic activities are
minimal in the park, limited primarily to bordering areas, the visitor center, and a few roads. Many of the sample ponds are clustered around the roads running laterally through the park. Because resistance caused by vegetation cannot be completely isolated from resistance caused by distance, the close proximity of the ponds around the roads may have led to the selection of human influence as the least resistant habitat type. In this case, gene flow across roads would have been high because of the short distance between ponds and the relatively small sizes of the roads, not necessarily because asphalt is a desirable migration medium for amphibians. Additionally, the roads running through the park experience low traffic, which is likely even further reduced at night when *A. opacum* movements typically occur. While prior studies have consistently found that roads create significant resistance to amphibian movements, it is difficult to separate the effects of roads and associated human development, such as fences and buildings (Mader 1984; Gibbs 1998; Carr and Fahrig 2001; Parris 2006; Hamer and McDonnel 2008). The occurrence of these associated structures at Mammoth Cave is minimal. Potentially, roads independent of other development may provide low-resistance pathways for dispersal. On rainy nights or at any time surface water remains on the roads, salamanders may be able to move easily along the roads, with little threat from oncoming traffic. However, I cannot make a firm conclusion without direct evidence from the roads at MCNP.

Water habitat, assigned the second lowest resistance value, consists primarily of the Green River. *Ambystoma opacum* adults are terrestrial, making movement across the river unexpected. Additionally, crossing the river would expose the salamanders to predatory fish, potentially selecting against such movements (Hecnar and M’Closkey 1997). Individuals may be drawn to the river by surrounding wet deciduous habitat, and
slow currents may facilitate movement across the river from one portion of suitable habitat to another. This differs from other models developed to describe movements of the *Ambystoma* genus, where moving water was assigned a resistance value higher than forest (Compton *et al.* 2007). However, this model was developed based on expert opinion, and one of the advantages of ResistanceGA is its freedom from any potential bias from expert opinion. If the Green River is facilitating movement within the park, the current would carry salamanders downstream. In this case, gene flow would likely occur only in that direction. The direction of gene flow could be evaluated in further analysis of the genetic data.

The resistances of forest types are consistent with another study in which *A. opacum* preferred migration paths with higher substrate moisture (Jenkins 2006). Moisture is an important variable in habitat selection, as amphibians experience high water loss through the skin (Spight 1968; Rothermal and Semlitsch 2002). Prolonged movement through drier areas may lead to desiccation, making these habitat types less conducive to migration between breeding ponds. Thus, we would expect the resulting resistances, where wet deciduous forest is the least resistant to movement, followed by coniferous and dry deciduous forest types.

While this analysis suggests human influence on this landscape does not restrict movement between breeding ponds, we are not suggesting that conservation methods involve the construction of more roads in the park. The low resistance in this model is likely a function of the low overall area occupied by areas of human influence and the proximity of many of the ponds along the roads. Roads may pose higher resistance in accumulation between breeding sites than single roads between a close pair of ponds.
Efforts to maintain natural habitat, particularly wet deciduous forest, between breeding ponds should be the focus of conservation of *A. opacum* at MCNP. The resistance of roads in this habitat should not be generalized, as the low traffic conditions differ from those of many other road types.

The use of stochastic genetic algorithms in the optimization creates potential variability in the output. Because of this, optimizations should be run at least twice. A repeat optimization is running on the WKU Bioinformatics Center’s supercomputer to confirm model selection. Future analyses of the landscape genetics of *A. opacum* at Mammoth Cave could involve optimization of other variables, such as slope or canopy cover, or analysis of the direction of gene flow to consider the potential effects of current in the Green River.
REFERENCES


Martin, JC. 2013. Landscape genetics of the marbled salamander (Ambystoma opacum) at Mammoth Cave National Park. Western Kentucky University.


Figure 1. Sample ponds 2 (top left), 14 (top right), and 45 (bottom).
**Figure 2.** Plot of GIS input data. Red represents human influence, blue is water, light green is dry deciduous forest, green is wet deciduous forest, and dark green is coniferous forest. White points represent pond sample sites.
Figure 3. Plot of GIS output data. Areas with darker green have the highest resistance, and lighter areas have the lowest resistance. White points represent pond sample sites.
**Figure 4.** Heat map generated from Figure 3 showing gene flow of *A. opacum* in MCNP. Lighter areas represent low resistance and higher gene flow, and darker areas represent high resistance and reduced gene flow.
Table 1. Descriptions of habitat types (from Martin 2013).

<table>
<thead>
<tr>
<th>Habitat Category</th>
<th>Description</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet Deciduous Forest (WDF)</td>
<td>Successional Tuliptree Forest</td>
<td>Rich Appalachian Red Oak – Sugar Maple Forest</td>
</tr>
<tr>
<td></td>
<td>Successional Black Walnut Forest</td>
<td>Central Interior Beech – White Oak Forest</td>
</tr>
<tr>
<td></td>
<td>Beech – Maple Unglaciated Forest</td>
<td>Shumard Oak</td>
</tr>
<tr>
<td></td>
<td>Sycamore – Silver Maple Calcareous Floodplain Forest</td>
<td>Chinquapin Oak Mesic Limestone Forest</td>
</tr>
<tr>
<td></td>
<td>Rich Levee Mixed Hardwood</td>
<td>Pin Oak Mixed Hardwood</td>
</tr>
<tr>
<td></td>
<td>Bottomland Forest</td>
<td>Depression Forest</td>
</tr>
<tr>
<td></td>
<td>Southeastern Successional Black Cherry Forest</td>
<td>Sinkhole Pond Marsh</td>
</tr>
<tr>
<td></td>
<td>Successional Tuliptree Forest (Circumneutral Type)</td>
<td>Southern Cattail Marsh</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Buttonbush Sinkhole Pond Swamp</td>
</tr>
<tr>
<td>Dry Deciduous Forest (DDF)</td>
<td>Interior Low Plateau Chestnut Oak – Mixed Oak Forest</td>
<td>Nashville Basin Shingle Oak – Shumard Oak – Chinquapin Oak Forest</td>
</tr>
<tr>
<td></td>
<td>Interior Dry – Mesic White Oak – Hickory Forest</td>
<td>Southern Red Oak Flatwoods Forest</td>
</tr>
<tr>
<td></td>
<td>Chinquapin Oak Unglaciated Bluff Woodland</td>
<td>Southern Red Oak – Mixed Oak Forest</td>
</tr>
<tr>
<td></td>
<td>Western Highland Rim Post Oak Barrens</td>
<td>Interior Low Plateau Chestnut Oak Forest</td>
</tr>
<tr>
<td></td>
<td>White Oak – Mixed Oak Dry-Mesic Alkaline Forest</td>
<td></td>
</tr>
<tr>
<td>Human Influence (HI)</td>
<td>Water – Willow Rock Bar and Shore</td>
<td>Commercial</td>
</tr>
<tr>
<td></td>
<td>Highland Rim Limestone Cliff/Talus Rock</td>
<td>Human Influence</td>
</tr>
<tr>
<td></td>
<td>Seep</td>
<td>Parking Lot</td>
</tr>
<tr>
<td></td>
<td>Rock</td>
<td>Road</td>
</tr>
<tr>
<td></td>
<td>Soil</td>
<td>Residential</td>
</tr>
<tr>
<td></td>
<td>Agriculture</td>
<td>Successional Broomsedge</td>
</tr>
<tr>
<td></td>
<td>Lawn</td>
<td>Vegetation</td>
</tr>
<tr>
<td></td>
<td>Power Line Easement</td>
<td>Cultivated Meadow</td>
</tr>
<tr>
<td></td>
<td>Building</td>
<td>Blackberry – Greenbrier</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Successional Shrubland Thicket</td>
</tr>
<tr>
<td>Water</td>
<td>Water</td>
<td>Naval</td>
</tr>
<tr>
<td>Coniferous Forest (CF)</td>
<td>Eastern Red-Cedar Successional Forest</td>
<td>East Central Hemlock Hardwood Forest</td>
</tr>
<tr>
<td></td>
<td>Early-Successional Shortleaf Pine Forest</td>
<td>Virginia Pine Successional Forest</td>
</tr>
<tr>
<td></td>
<td>Appalachian Low-Elevation Mixed Pine/Hillside Blueberry Forest</td>
<td>Virginia Pine – Red-Cedar</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Successional Forest</td>
</tr>
</tbody>
</table>
**Table 2.** Description of models considered in analysis. FST represents pairwise Fst values, GEO and LAND are fixed effect terms representing pairwise geographic distances and vegetation types, respectively. PPID is the random effect term of identifiers for each unique pairwise pond combination.

<table>
<thead>
<tr>
<th>Name</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null Model</td>
<td>FST ~ 1 + (PPID)</td>
</tr>
<tr>
<td>Distance (IBD Model)</td>
<td>FST ~ GEO + (PPID)</td>
</tr>
<tr>
<td>Vegetation (IBR Model)</td>
<td>FST ~ GEO + LAND + (PPID)</td>
</tr>
</tbody>
</table>
Table 3. Objective function table used in model selection. LL is the log likelihood value, AIC is the Akaike Information Criterion, and AICc is the corrected value.

<table>
<thead>
<tr>
<th></th>
<th>LL</th>
<th>AIC</th>
<th>AICc</th>
<th>Marginal $R^2$</th>
<th>Complete $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Null</td>
<td>2981.5680</td>
<td>-5957.1360</td>
<td>-5961.0526</td>
<td>0.0000</td>
<td>0.3118</td>
</tr>
<tr>
<td>Distance</td>
<td>2995.4689</td>
<td>-5982.9377</td>
<td>-5986.6824</td>
<td>0.0405</td>
<td>0.3348</td>
</tr>
<tr>
<td>Vegetation</td>
<td>2998.6543</td>
<td>-5989.3086</td>
<td>-5983.3551</td>
<td>0.0540</td>
<td>0.3587</td>
</tr>
</tbody>
</table>
Table 4. Resistance values assigned to each habitat type.

<table>
<thead>
<tr>
<th>Habitat Category</th>
<th>Resistance Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WDF</td>
<td>1.8279</td>
</tr>
<tr>
<td>DDF</td>
<td>6.5647</td>
</tr>
<tr>
<td>HI</td>
<td>1.0000</td>
</tr>
<tr>
<td>W</td>
<td>1.6207</td>
</tr>
<tr>
<td>CF</td>
<td>5.8722</td>
</tr>
</tbody>
</table>