Rodent Population Connectivity in Coffee Agroecosystems

by

Beatriz Otero Jiménez

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in the University of Michigan
2019

Doctoral Committee:

Professor Priscilla K. Tucker, Co-Chair
Professor John H. Vandermeer, Co-Chair
Professor Timothy James
Professor Ivette Perfecto

Beatriz Otero Jiménez

botero@umich.edu

ORCID iD: 0000-0002-9686-9493

Dedication

To *mami* and *papi* for their unconditional support

Acknowledgements

I would first like to thank my parents, for encouraging me to pursue my interest in the natural world, without them and their sacrifice I would not be here. I would like to thank Carlos J. Anderson for his support during the dissertation process and crucial support with statistical and computational analysis. I thank my co-advisor Priscilla Tucker for all her support and patience in the past 8 years of my graduate career. Thank you for being available and willing to talk, read, and review all my work. I especially want to thank her for her support of my outreach and education interests while completing my doctoral work. I want to thank John Vandermeer for his excitement for science, thoughtful questions, and support in pursuing my research project. I would also like to thank John Vandermeer and Ivette Perfecto for having a space for social and political consciousness in their research and within their laboratories, knowing there is space for these conversations and actions in academia has encouraged me to continue in this career path.

I would like to thank El Colegio de la Frontera Sur in San Cristobal, Chiapas Mexico, especially Dr. Consuelo Lorenzo and Jorge Bolaños for their support in sample collection and in coordination and execution of the El Triunfo Bisphere Reserve sampling. Thanks to Dr. M. Raquel Marchan Rivadeneira for assistance with genetic analysis and her support above and beyond her call of duty. Thank you to the managers, farmers and owners of Finca Irlanda and Finca Hamburgo in Chiapas, Mexico, and the park rangers and employees at the El Triunfo Biosphere Reserve for allowing us to conduct this study and for their support with fieldwork.

Table of Contents

List of Tables

List of Figures

Abstract

Humans have been modifying the Earth's land surface for millennia. In the last 300 years these changes have increase in intensity and spatial extent. In tropical regions these anthropogenic changes are dominated by the expansion of agriculture. This has led the majority of remaining tropical forests to exist as fragments embedded in a matrix of agricultural production. This production varies in type and diversity of crops and management practices, through which organisms must navigating through or surviving within the matrix. For this reason, understanding the effects of these agricultural landscapes on species dispersal is crucial in the development of successful conservation planning.

The purpose of this dissertation was to explore the effects of varying coffee production management practices on the population structure and connectivity of tropical rodents. This study was conducted in the coffee growing region of Soconusco and the El Triunfo Biosphere Reserve in Chiapas, Mexico. We used genetic and landscape data to study the population structure and connectivity of two common rodent species, *Heteromys desmarestianus goldmani* and *Peromysucs gymnotis*, in this landscape.

We found that levels of population connectivity and genetic diversity vary between the two sampled species, which is supported by their differences in ecological specialization. *Heteromys* in the coffee farms were characterized by subtle genetic structure, which correlates with high management intensity coffee production and high genetic diversity. On the other hand, *P. gymnotis* individuals showed no signal of population structure and lower degrees of genetic diversity. When comparing *H. d. goldmani* populations from the continuous forest (El Triunfo Biosphere Reserve) and the coffee production region we found similar levels of genetic diversity,

suggesting that high levels of migration and gene flow can be maintained in the coffee agroecosystem.

This study highlights the potential of integrating molecular and landscape data to explore population connectivity of elusive species, such as terrestrial small mammals. It also shows the importance of studying the responses to environmental change for species with different levels of ecological specialization within a group, since these responses can vary. Additionally, it identifies coffee production as an important refuge for rodent species within anthropogenic landscapes. This work adds to the growing body of literature in landscape genetics by demonstrating that rodents can show population structure at small scales and that this structure can be driven by landscape factors linked to agricultural management.

Chapter 1 Landscape Drivers Of Connectivity For A Forest Rodent In A Coffee Agroecosystem

1 Introduction

The majority of landscapes are characterized by a highly heterogeneous structure. This structure is driven by natural (e.g., humidity, topography) or anthropogenic forces (e.g., agriculture, urbanization). From the point of view of an organism, some landscape characteristics will support the organism's survival and reproduction, while others will restrict it. Areas that promote the survival and reproduction of organisms are known as habitat, while the matrix is the land in between habitable areas for a particular species, which usually encompasses most of the landscape (Perfecto et al. 2010). The metapopulation framework introduced by Levins (1969) helps us in the study of populations inhabiting these heterogeneous landscapes.

Metapopulations consist of several subpopulations of an organism separated in space and with some probability of survival and migration from one subpopulation to another (Levins 1969, Haski 1999). This movement can also be referred to as population connectivity. Maintaining high levels of population connectivity can allow individuals to locate new resources and prevent inbreeding depression (Olah et al. 2017). Among the many forces influencing dispersal is the quality of the landscape matrix (Vandermeer and Carvajal 2001). The matrix can vary in quality depending on the species in question and the land cover represented in the area (Perfecto et al. 2010). High quality matrices will promote species movement across the

landscape, while low-quality matrices will impede movement. Understanding the influence of matrix composition on population connectivity is of great importance in our current world where the landscape is increasingly fragmented due to human perturbations, such as agriculture which accounts for approximately half of the Earth's land surface (Hooke et al. 2012). As a result, most tropical forests exist as fragments embedded in a matrix of agricultural production. Frequently, agriculture is viewed in opposition to biodiversity conservation (Perfecto and Vandermeer 2008).

However, within the agricultural matrix we can find management practices that vary greatly in the level of intensification. Agricultural intensification is the transition from high crop diversity and low external inputs (e.g., organic or traditional production) to systems with a single crop species and increased use of external inputs (i.e., pesticides, herbicides; Perfecto et al. 2010). Because of this variation in management, we do not expect all agricultural lands to influence species movement equally. Research has highlighted the important role traditional agriculture and agroecological production practices can play in the conservation of many taxa (Vandermeer and Perfecto 2007, Philpott et al. 2008) while also providing sustainable livelihoods for farmers (Altieri 1999, Perfecto and Vandermeer 2010). Few studies evaluate the effect of different management practices on species dispersal at a fine scale (for an example see Flores-Manzanero et al. 2018). By identifying which agricultural practices facilitate dispersal we can maintain productive landscapes while also preventing species extinctions. The agricultural matrix is an important part of the conservation landscape that can provide habitat for species and dispersal corridors between forested areas.

Studying the influence of agricultural practices on species requires measures of population connectivity. Direct measures of connectivity are difficult to obtain. They require following an individual's movement through the landscape using mark-recapture, or telemetry

methods. While these methods provide detailed information on the movement of individuals, sample sizes are invariably limited and only a static snapshot of the movement is obtained, as recorded during the time of the study. Additionally, small and elusive species are difficult to include in this type of study.

Landscape genetics, a field that incorporates population genetics, landscape ecology and spatial statistics to study the effect of landscape characteristics on populations provides an alternative to direct measures of population connectivity by using genetic distance (i.e., gene flow) between individuals or populations as a proxy for average dispersal and provides a measure of population connectivity that can go back several generations depending on the methods used. Using this genetic information, we can identify landscape variables that drive genetic patterns by methods such as resistance surface modelling (Spear et al. 2010). Landscape resistance models are created for each landscape variable of interest and these can be compared to the measures of genetic distance to determine which variables better explain the observed genetic pattern. The results are assumed to reveal which variables are promoting or impeding movement.

This study explores the way in which different landscape characteristics result in genetic correlations among individuals and subpopulations of a small forest dwelling rodent, *Heteromys desmarestianus goldmani,* in order to understand how matrix quality, including connectivity, influences genetic structure in a landscape dominated by the coffee agroecosystem. Rodents are the most diverse group of mammals, playing a variety of ecological roles, such as seed dispersers, herbivores, predators, prey, disease vectors, and ecosystem engineers, both positive and negative. As members of diverse food webs (Dickman 1999) they are frequently thought to be key elements of ecosystems (Brown and Heske 1990, Davidson and Lightfoot 2006). Despite

this acknowledged importance, rodents are fairly understudied, especially in tropical systems (Chung and Corlett 2006). Although literature exists on general natural history (e.g., Fleming 1983, Quintero and Sanchez-Cordero 1989), only a few studies are concerned with the influence of landscape features on these common species (but see Flores-Manzanero et al. 2018, Otero-Jiménez et al. 2018). Recently, with the development of new molecular techniques, genetic data has been used to infer dispersal, facilitating the study of smaller species, such as rodents, for which data collection using direct methods (mark and recapture or telemetry) is more difficult.

The spiny pocket mouse (*Heteromys. desmarestianus goldmani*) is a common forest dwelling rodent in southern Mexico and Central America. *H. d. goldmani* is a seed predator and important seed disperser in tropical forests (Fleming 1983). Despite being a commonly found species, we know very little about its population dynamics and dispersal. We collected population genetic data from microsatellite markers and land cover data derived from satellite images to identify potential landscape drivers influencing population connectivity. We expect physical and structural landscape variables, some of which are related to agricultural management, to have a significant effect on population connectivity. More specifically, we expect proximity to rivers or streams (i.e., riparian effect) and high canopy cover to promote dispersal of *H. d. goldmani*. These characteristics are likely to provide important resources needed to survive within the farms as well as protection from predators. They can also help to maintain temperature and humidity levels.

2 Methods

2. 1 Study Species

The spiny pocket mouse, *H. desmarestianus*, is a rodent that inhabits moist forest habitats from southern Mexico to Panama (Fleming 1974). It is the most abundant small mammal in these

regions (Klinger 2007) and has a home range of 100 m^2 (Fleming 1974), which is small for a terrestrial rodent. Studies conducted in Costa Rica have shown that the diet of *H. desmarestianus* mainly consists of palm nuts and other seeds (Fleming 1983), making this rodent an important seed disperser in tropical regions (DeMattia et al. 2004). Based on genetic and morphological data *H. desmarestianus* is considered a species complex of which, *H. d. goldmani* is a member (Rogers and Gonzalez 2010). Little is known about the population structure and dispersal pattern of this common species.

2.2 Study Site

Coffee production in Latin America represents an ideal system for studying the effects of different management practices. In this region, coffee is produced in a variety of ways that follow an intensification gradient (Fig. 1-1a). It is common to find several of these management practices represented in small areas. For this reason, we chose to do our study in the tropical montane region of Soconusco in Chiapas, Mexico (Fig. 1-1b). This area is dominated by coffee production that varies in management intensity, ranging from rustic to unshaded monocultures with forest patches scattered between them (Fig. 1-1c). Farms included in the study have been producing coffee for 60 to 100 years. Although management practices vary over time, these farms have had similar management practices for at least the past 20 years (Perfecto and Vandermeer 2002).

2.3 Field Sampling

H. d. goldmani samples were collected from 2012 to 2015. We used 136 samples previously collected in the region from 2012 to 2014 (Otero Jiménez et al. 2018) and added 29 new samples from 2015 with the goal of reaching continuous sampling across the study site. Ear tissue samples from *H. d. goldmani* were collected from six sites: three forest fragments and

three coffee farms of various management levels that are adjacent to the forest fragments (Fig.1-1c). Coffee farms were categorized by the level of management intensity (i.e., Low, Medium, and High) based on the Moguel and Toledo (1999) classification system (Fig. 1-1a). To facilitate continuous animal collection across this complex landscape we sampled an area of approximately 4 km x 2 km (Fig. 1-1c), by following roads and trails within the farms and the forest fragments, following the methodology of Otero Jimenez et al. 2018. Ear tissue samples were preserved in 20% DMSO buffer saturated with NaCl. Mice were captured using 22.9 cm x 7.6 cm x 8.9 cm Sherman live traps. Sex and GPS coordinates for each individual sample were recorded. Animals were handled in accordance with the University of Michigan's Committee on Use and Care of Animals.

2.4 Genetic Data

DNA Extraction – In order to obtain genetic population data we extracted DNA from 165 adult and juvenile *H. d. goldmani* ear tissue samples within our sampling area. Genetic data were obtained using the 12 species specific microsatellite markers developed by Otero-Jiménez et al. 2018 (HET1, HET4, HET23, HET27, HET32, HET34, HET37, HET42, HET41, HET46, HET56, HET57). All loci were polymorphic in all sampling locations. Loci were tested for the presence of scoring errors and null alleles using Micro-Checker (Van Oosterhout et al. 2004). We tested for deviations from Hardy-Weinberg equilibrium (HWE) and for linkage disequilibrium at all loci and collecting sites (e.g., coffee farms of different intensities and forest fragment) using Arlequin 3.5.1.3 (Excoffier and Lischer 2010). Bonferroni corrections were applied to determine significance of HWE and linkage results. We estimated relatedness (r) within and between sites to check that individuals sampled were not siblings. We calculated pair-wise values of r for all individuals using GenAlEx (Peakall and Smouse 2012).

Isolation by Distance (IBD) – To assess the influence of geographic distance on the observed genetic structure we calculated isolation by distance (IBD). We measured the correlation between pairwise genetic distance of all sampled individuals using the individual measure of differentiation a_r (Rousset 2000), and Euclidian distance using a simple Mantel test in Arlequin (Excoffier and Lischer 2010). In addition, we estimated a Mantel correlogram based on 100 m distance classes considering the home range reported for *Heteromys* (Fleming 1974). Significance was calculated using Spearman correlation based on 10,000 permutations. Eucledian distances were estimated in Genalex (Peakall and Smouse 2012) and genetic distances (a_r) were estimated using Genepop (Raymond and Rousset 1995, Rousset 2008).

Population Structure Analyses – The coffee agroecological landscape arrangement was expected to have some effect on the dispersal of the mice (Otero Jiménez et al. 2018). We estimated the number of genetic units (K) and the locations of breaks in gene flow that define these clusters using Geneland (Guillot et al. 2008), a Bayesian clustering method. The use of Bayesian clustering removes the limitation of a priori population assignment; it identifies genetic units based on multilocus genotype data by maximizing HWE and minimizing linkage disequilibrium (Manel et al. 2005; Latch et al. 2006). We used Geneland 4.0.3 (Guillot et al. 2005) to calculate the number of genetic units in our study. We used Geneland in this analysis because it has been shown to detect weak genetic structure in areas with gene flow (Safner et al. 2011).

The Geneland analysis included 20 independent runs with 10 000 000 MCMC iterations and 10 000 thinning (i.e., saving results from one iteration every 10 000), while varying K from 1 to 10. Correlated and null allele model options were activated and the potential error for spatial

coordinates was set at 10 m. We used default settings for all other parameters. Optimal K was inferred from the run with the greatest average likelihood. After determining the optimal number of subpopulations (K), a separate run was performed for the assignment of individuals. For these runs, K was set to the previously inferred optimal number of subpopulations, the run parameters were 5 000 000 MCMC and 5 000 thinning. We calculated the posterior probability of subpopulation membership for each pixel of the spatial domain (500 x 500 pixels) with a burn-in of 100 for the run with the highest posterior probability.

For the genetic clusters identified by Geneland, we measured genetic diversity by quantifying observed heterozygosity (HO), expected heterozygosity (HE), fixation index (FIS) and allelic richness (AR) using GenAlEx (Peakall and Smouse 2012). To assess differences in genetic diversity between clusters we used values for each measure (i.e., F_{IS}, H_E, H_O, AR) for each locus at each cluster. We conducted a bootstrapping analysis of the mean in R. To assess genetic differentiation between groups we calculated pairwise F_{ST} (Wright 1951) using Arlequin (Excoffier and Lischer 2010).

2.5 Landscape Data

Landscape variables – Within a 4 km long (E-W) by 2 km wide area (Fig. 1-1c), we tested the significance of 5 landscape features we hypothesized could influence resistance to dispersal: 1) tree cover (TC), 2) slope (S), 3) elevation (E), 4) riparian effect (RE), and 5) streams (STR; Table 1).

Landscape features were quantified using products derived from aerial imagery (Google 2015) and a 20 m resolution digital elevation model (DEM). Land cover was digitized manually using the heads-up digitizing method (Bolstad 2016) to map polygons of similar levels of tree cover discernable at a 1:2,000 scale. Land cover was classified into a ranking of increasing

apparent habitat quality: 1) no high vegetation cover (open), 2) sun coffee plantation (no tree cover), 3) 25% tree cover coffee plantation, 4) 50% tree cover coffee plantation, 5) 75% tree cover coffee plantation, and 6) 100% tree cover (Fig. 1-2). Slope and elevation layers were developed from the DEM using appropriate geoprocessing tools (Fig. 1-2). Streams were delineated from the DEM by calculating flow accumulation and classifying any cells above a value of 100 as a stream (Fig. 1-2). We created the riparian effect layer by calculating a Euclidean distance raster from stream locations, which we used to represent lower resistance with closer proximity to a stream (Fig. 1-2). Digitization and DEM raster processing was performed with ArcGIS 10.3.1 (ESRI, 2015).

Resistance Surfaces – Each landscape variable layer was converted to a raster map representing the assumed resistance to movement of each variable to *H. d. goldmani* (Table 1-1, Fig. 1-2). These predictions of landscape effects on movement were based on the available natural history research of *H. d. goldmani* and other members of the *Heteromys* genus (Fleming 1974, Fleming 1984, Martínez-Gallardo and Sánchez-Cordero 1993, Klinger 2007). Landscape surfaces were rescaled to the finest possible resolution (20 m x 20 m) to reflect the hypothetical perceived scale of resistance for *H. d. goldmani.*

2.6 Landscape Genetic Analysis

Most studies evaluating the effects of landscape resistance on genetic population connectivity rely on expert opinion to describe the relationship between landscape variables and species movement (Zeller et al. 2012). We used the method developed by Peterman et al. 2014 to optimize our resistance surfaces to address concerns of analysis based on expert opinion. For example, many studies using expert opinion to develop landscape resistance models treat expert

opinion as empirical data, when it is not, thus making it difficult to evaluate performance (Zeller et al. 2012). In general, it is difficult to describe accurately the ecological processes being modelled in resistance analyses, and even if the processes are known, there is no guarantee that they will have significant influence on gene flow (Peterman et al. 2018). Peterman et al.'s (2014) optimization approach relies on a genetic algorithm to explore the parameter space, which seeks to maximize the relationship between pairwise landscape resistance distances (i.e., landscape variables) and pairwise genetic distances without any a priori assumptions (Peterman et al. 2018).

The optimization was done using the R package ResistanceGA (Peterman et al. 2018). Landscape surfaces were optimized in two steps: (1) single resistance surface optimization, and (2) combined (i.e., multi-) surface optimization (Fig. 1-2).

For the first step we optimized each landscape surface independently, using the *commuteDistance* function exploring resistance values up to 2,500 using an 8 neighbor joining scheme to measure connectivity (Fig. 1b). The *commuteDistance* function, which calculates the commute distance or the time it takes for an individual to move from point *a* to point *b* in a particular landscape (vanEtten 2017). Our dependent variable was genetic distance measured as a_r (Rousset 2000) and our predictor variables were the landscape resistance values. For continuous landscape variables (i.e., slope, elevation and riparian effect) we tested 9 possible transformations for the resistance relationship between both variables (i.e., linear, monomolecular, reverse monomolecular, inverse monomolecular, inverse-reverse monomolecular, Ricker, reverse Ricker, inverse Ricker, inverse-reverse Ricker). Additionally, we included pair-wise geographic distance of samples as our null model. We conducted 3 independent optimization runs for each of the landscape variables to ensure results were robust.

The optimal resistance surface was identified using the Akaike Information Criterion (AIC), determined by a linear mixed effects model with a maximum likelihood population effects parameterization (MLPE; Peterman et al. 2018). We then conducted a bootstrap analysis in which 75% of the samples were randomly selected without replacement and each surface was fit to the subset of samples. For each subset of samples, the models' average rank, average weight and percentage that a surface was selected (top rank) were calculated with 10 000 iterations. Before we continued to the second optimization step for combined surfaces, we conducted a Spearman coefficient correlation test in R between the top commute distance matrices of all optimized single surfaces that performed better than geographic distance alone. We selected variables that showed small to moderate correlation ($p < 0.49$), to avoid including correlated variables in the combined surface model.

For the second optimization step we performed the combined surface optimization with surfaces that showed moderate to low correlation (Fig. 1-2). Parameters for the combined surface optimization were the same as for the first optimization step. Additionally, we performed bootstrap model selection using the same parameters as for the single surface optimization in order to obtain the average rank, average model weight and the top ranked model for individual and combined surfaces. Optimization R script was adapted from Flores-Manzareno et al. 2018 to conduct this analysis.

3 Results

3.1 Microsatellite Analysis

We found evidence of linkage disequilibrium at one locus (HET-41) across multiple sampling sites. This locus was eliminated from the data set for all further analyses. Micro-Checker found no evidence of scoring errors but did imply null alleles for locus HET-46 (P <

0.0001). All loci were polymorphic in all sampling locations. A total of 97 alleles were scored at 11 loci in all *H. d. goldmani* samples with an average of 7 alleles per locus (range 3 – 13; Table A1-1). Primer and locus information can be found in Otero Jiménez et al. 2018. Results for the Lynch and Ritland (1999) relatedness estimator show values of $r < 0.5$ within and between sites (Table A1-2) indicating samples were not taken from members of the same litter.

3.2 Population Structure

Results from genetic clustering analysis revealed 6 distinct genetic clusters (Fig. 1-3; Table A1-3). Individuals from each of the 3 forest fragments sampled were classified as distinct populations (Clusters 2, 4 and 5; Fig. 1-3). Individuals sampled in the low intensity coffee farm were clustered with individuals from the adjacent forest fragment (Cluster 4; Fig. 1-3). Individuals from the medium and high intensity coffee farms show a similar trend, i.e. individuals from localities close to the forest fragment cluster with forest individuals (Cluster 4) whereas individuals from localities further east or west of the forest fragment, constitute separate clusters (Clusters 1 and 3, Fig. 1-3). Additionally, we found two individuals from the high intensity coffee farm that were assigned to the medium intensity farm cluster (Cluster 1, Fig. 1-3), suggesting that these might be early generation migrants. We also found a single individual assigned to its own cluster (Cluster 6). This individual could be a member of a different population to the South of the study site.

For the genetic diversity and cluster differentiation analysis we removed Cluster 6, because it only has one individual. For the rest of the 5 Clusters we did not find any significant difference in genetic diversity measures between the clusters (i.e., F_{IS}, H_E, H_O, AR; Fig. A1-6). Pair wise genetic differentiation (F_{ST}) between clusters was statistically significant with values ranging from 0.01 to 0.05 (Table 1-2).

3.3 Isolation by Distance

Isolation by distance results show significant but weak positive correlation between genetic distance and geographic distance for all samples (Mantel r = 0.11, P = 0.002; Fig. A1-7). The mantel correlogram showed significant positive relationship between genetic and geographic distance for 4 distance classes < 500 m (Fig. A1-8; Table A1-4). Additionally, it identified 5 distance classes with a weak but significant negative relationship (Fig. A1-8; Table A1-4), which we interpret as having no biological significance for the species.

3.4 Landscape Genetic Analysis

Model selection results show that the slope (S) resistance surface was the best-supported model (47.2% of the times based on 10,000 bootstrap replicates; Table 1-3). From all the transformations tested, Inverse-Reverse Ricker function was the best fit for our genetic distance data (Fig. 1-4a). The next best supported resistance surface was tree cover (TC) (38.2% of the times; Table 1-3). Additionally, the linear mixed effect model identified both variables (i.e., slope and cover) as significant predictors of genetic distance. Streams and elevation explained 12.4% and 1.4% respectively, of the variation in genetic distance data than distance alone (Table 1-3). Riparian effect had poor performance, being the only landscape model that explained less than distance alone (Table 1-3).

We tested for correlations between all layers, except riparian effect, using Spearman's tau. All surfaces showed some degree of correlation between each other. Seven out of the 9 showed strong (> 0.50 Spearman tau; Table A1-5) and 3 layer pairs showed evidence of moderate correlation (S + TC, S + STR, S + RE; Table A1-5). For the next optimization step, we chose to perform a combined surface optimization analysis with a single combined surface of slope and tree cover, along with all other single surface variables. Because tree cover (TC),

streams (STR) and riparian effect (RE) had strong correlations with each other we only created a single combined surface. The lower correlation between tree cover and slope could be explained by the differences in management practices between the medium and high intensity coffee farms. Steeper areas in the medium intensity coffee farm are more likely to be left unmanaged, while in the high intensity farm steeper areas are included in coffee production. The slope variable had a mean contribution to the model of 12.4% and tree cover of 9.8% (Table 1-4). Results showed the combined surface to be the best supported model at 69.1%, followed by slope and tree cover (Table 1-4).

4 Discussion

In this study we set out to investigate the effect of landscape variables on the population connectivity of the forest rodent *H. d. goldmani* in a coffee agroecosystem. The study region included different farms with varying management practices that range from low to high intensity of coffee production. A previous study found limited population structure in this coffee growing region for *H. d. goldmani* with patterns of gene flow not easily explained by geographic distance (Otero Jiménez et al. 2018). Given the small home range of *H. d. goldmani* (100 m^2) compared to other rodent species (Fleming 1974), it is not surprising to find genetic structure at a small spatial scale as observed in Otero Jiménez et al. (2018). This study was designed to search for potential relationships between landscape features and the population connectivity patterns observed in Otero Jiménez et al. (2018). Study results showed higher population connectivity between individuals sampled near and inside the central forest fragment, while connectivity between individuals sampled within the coffee plantation decreased with distance from the forest, generating separate genetic populations and suggesting an influence of management practices on connectivity (Otero Jiménez et al. 2018).

In our efforts to detect the variables influencing the observed genetic structure and limited gene flow, we found a weak but significant increase in genetic differentiation with increasing geographic distance between individuals following Otero Jimenez et al. 2018. Previous studies have found a range of IBD responses in terrestrial rodent populations. Many studies have found weak or no relationship between genetic and geographic distance between individuals or populations (Chiappero et al. 2011, Gerlach and Musolf 2001, Flores-Manzanero et al. 2018), while others have found significant IBD (Nicolas et al. 2008, Berkman et al. 2018). However, these results are influenced by many variables including spatial scale of the study and dispersal capabilities of the studied species. It is expected that species with limited dispersal capability will show stronger signals of IBD (Wright 1943).

We then moved to examine other potential landscape variables that could be influencing these patterns. We generated resistances surfaces for 5 landscape variables (i.e., slope, elevation, streams, riparian effect and tree cover). We optimized the resistance surfaces of each using genetic distance information. These results showed that, from the variables sampled, slope is the one with the strongest influence on genetic distance, followed by tree cover. Not surprisingly, a combined surface of the two variables was the best model to explain the observed genetic structure. We found that our results are supported by studies done on other mammal species, where slope was an important predictor for richness and density (Russo et al. 2016, Carver 2010, Keeley et al. 2016). Additionally, tree cover or canopy cover has been identified as an important component to promote movement and species richness of terrestrial rodents (Caudill and Rice 2016, Santos-Filho et al. 2012, Lomolino and Perault 2001) and other groups (birds; Martensen et al. 2012, amphibians; Popescu and Hunter 2011).

As mentioned in the results, the permeability of areas with higher slopes can vary in the landscape with the level of management intensity. We found that intermediate slopes (i.e., 30 to 40% slope) had the lowest levels of resistance for *Heteromys* with increasing resistance at the two extremes. This could be due to the fact that flat (low slope) areas are used for coffee production and the mice are not able to move as easily through this landscape. However, as slope increases, harvesting and management of coffee is more difficult and the land is not as intensely managed, creating a better environment for the mice until it reaches a certain threshold where it becomes too steep for mice to move readily. However, these patterns have to be further studied to identify the relationship between slope and other landscape variables that may influence connectivity.

The best model of the tree cover surface showed that intermediate levels of tree cover (50 - 75%) are the most permeable for *Heteromys*, followed by high tree cover (> 75%). Areas with 25% tree cover were classified with the highest resistance (Fig. A1-9). Surprisingly, open areas and sun coffee, which are the smallest percentage of cover type in the area (Fig. A1-9), were classified with low resistance values. This could be due to the absence of samples in those areas, which could make them irrelevant in the resistance calculations since no individuals would have connectivity paths that cross these regions. The Caudill and Rice (2016) study on rodent diversity in a similar coffee system highlighted the importance of canopy cover for rodent species richness and density. These results along with our findings suggest that tree cover is essential for maintaining rodent population connectivity in agroecosystems.

We expected riparian effect to be one of the main variables explaining the observed genetic structure. Our results showed the opposite trend, with this variable performing the worst from all those analyzed. Because this measure was based on flow accumulation calculated from a

digital elevation model (DEM), it might not adequately reflect the desired patterns of riparian effect or account for variation in management along streams across farms with different production practices. Instead, our results show the streams resistance surface as one of the top 3 models, where the top model showed streams to have low resistance to movement of *Heteromys,* suggesting the influence of streams on *Heteromys* population connectivity.

In summary, our results show that *H. d. goldmani* population connectivity and structure is influence by landscape variables, more than by geographic distance. This was shown by our IBD calculations and the resistance surface models. Our study had several limitations that could have influenced the strength of the patterns observed. One of these limitations was sampling; we worked within active coffee farms and in a highly mountainous region where certain areas were inaccessible for trapping, thus limiting our capacity for continuous sampling. In addition, sampling in areas with little to no tree cover resulted in no trap success. Secondly, our landscape data were limited in grain by the resolution of the satellite images and the digital elevation model used, which was 20 m x 20 m at its finest. Studies have demonstrated the importance of spatial scale in determining the relationship between landscape variables and population structure (Zeller et al. 2012, Oyler-McCance et al. 2013). Scaling of the resistance surfaces must reflect the scale at which the study species utilizes the environment. Since our project examined rodent populations a finer resolution would have been preferable, however, because 20 m is still smaller than the estimated home range of *Heteromys*, this analysis has provided some valuable insight. We hope that with technological advancements collecting these data at finer scales will become more affordable and accessible in the near future. We expect that with finer scale data we will find stronger relationships between the landscape variables analyzed and the genetic structure of *Heteromys*. Finally, while population structure and connectivity are influenced by many

variables, we limited our study to five of these variables and there are many others that could be included and potentially have a stronger influence. For example, humidity has been explored as an important variable in small mammal studies (Flores-Manzanero et al. 2018) and could be an important measure for *H. d. goldmani* because this is a forest dwelling rodent.

This study highlights the potential of integrating molecular and landscape data to explore population connectivity of elusive species, such as terrestrial small mammals. Tropical rodents play an important role in the ecosystem and their conservation is important for the maintenance of other species (Fischer et al. 2011). From the literature we know that some small mammals are being negatively impacted by fragmentation, land cover and habitat loss (Blois et al. 2010). Recent studies evaluating the responses of terrestrial rodents to land cover changes and fragmentation have shown that responses vary among species depending on their diet and habitat preferences (e.g., Fischer et al. 2011, Passamani and Fernandez 2010). However, studies evaluating the genetic structure and population connectivity of tropical small mammals are lacking in the scientific literature (but see Flores-Manzanero et al. 2018, Otero Jiménez et al. 2018, Balkenhol et al. 2013). Even rarer are studies that address the effects of the matrix composition on the connectivity of populations. Our study adds to this growing body of literature, showing that rodents can show population structure at small scales and that this structure can be driven by landscape factors linked to agricultural management.

5 Conclusions

This study provides an example of how genetic and landscape data can be used to study population connectivity in agricultural systems at a small scale. In addition, the findings add important insights to the literature on tropical rodents and their responses to anthropogenic land use changes. Our results highlight how agricultural production (e.g., tree cover) can be managed

to promote and even sustain rodent populations, thus emphasizing the potential role agriculture can play in maintaining biodiversity and connectivity in tropical landscapes.

6 Tables

Table 1-1 Expected *H. d. goldmani* resistance to movement for each landscape variable.

Landscape Variable	Variable Type	Expected Resistance
Euclidean Distance		Isolation by distance. Greater resistance when further apart
Elevation	Continuous	Greater resistance as you move away from the optimal elevation
Slope	Continuous	Greater resistance as the percent slope increases
Riparian Effect	Continuous	Greater resistance further from stream edge
Tree Cover	Categorical	Greater resistance with lower tree cover
Streams	Categorical	Streams as barriers to movement

Table 1-2 Results for genetic differentiation (FST) between genetic clusters identified by Geneland analysis. Cluster 6 was not included in this analysis because only a single sample was assigned to this cluster.

	Cluster	1 Medium-Intensity Coffee	2 East Forest Fragment	3 High-Intensity Coffee	4 Central Forest Fragment	5 West Forest Fragment
1	Medium-Intensity Coffee	0				
2	East Forest Fragment	0.0425	0			
3	High-Intensity Coffee	0.0208	0.0241	0		
4	Central Forest Fragment	0.0241	0.0461	0.0193	0	
5	West Forest Fragment	0.0291	0.0319	0.0196	0.0256	0

Table 1-3 Model selection results for the generalized linear mixed -effects models optimized on genetic distance (ar) for *H. d. goldmani*. K is the number of parameters for each model.

Surface	K	Equation	AIC	Average Weight	Average rank	Top Model %
		Inverse-Reverse				
Slope	4	Ricker	-13067	0.358	2.333	43.05
Tree Cover	7	NA	-13067.15	0.282	2.363	33.75
Streams	3	NA	-13064.23	0.133	3.520	11.71
Riparian Effect	4	Ricker	-13063.53	0.115	3.997	10.49
Elevation	4	Monomolecular	-13064.07	0.075	3.931	0.88
Distance	2	NA	-13062.35	0.035	4.857	0.12

Table 1-4 Model selection results for both individual and composite surfaces for *H. d. goldmani.*

Surface	K	Equation	AIC	Average Weight	Average rank	Top Model %	Variables
Combined	10	NA	-13072.51	0.353	1.971	45.51	Slope and Tree Cover
Slope	4	Inverse-Reverse Ricker	-13069.23	0.218	3.015	21.76	Slope
Tree Cover	7	NA	-13069.43	0.165	3.125	13.55	Cover
Streams	3	NA	-13066.4	0.098	4.370	9.23	Streams
Elevation	4	Monomolecular	-13066.32	0.049	4.851	0.57	Elevation
Riparian Effect	4	Ricker	-13065.68	0.090	4.874	9.33	Riparian Effect
Distance	2	NA	-13064.55	0.026	5.795	0.05	Distance

7 Figures

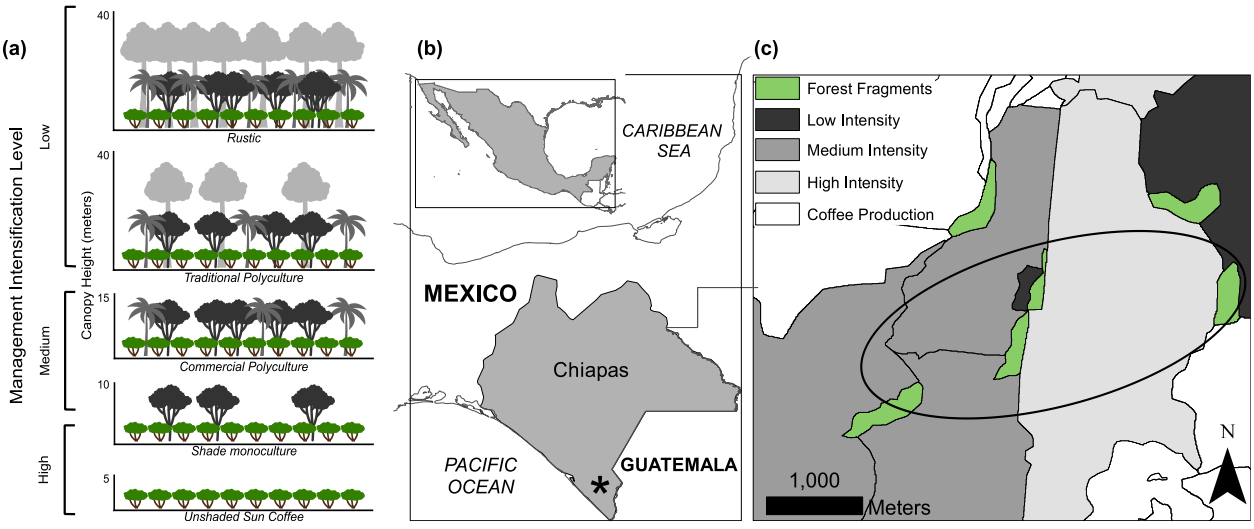


Figure 1-1 Coffee management intensification diagram and map of study area. (a) Diagram of coffee management intensification (based on Moguel and Toledo 1999) and (b, c) map of study region; b) shows the location of our study site within the state of Chiapas, MX

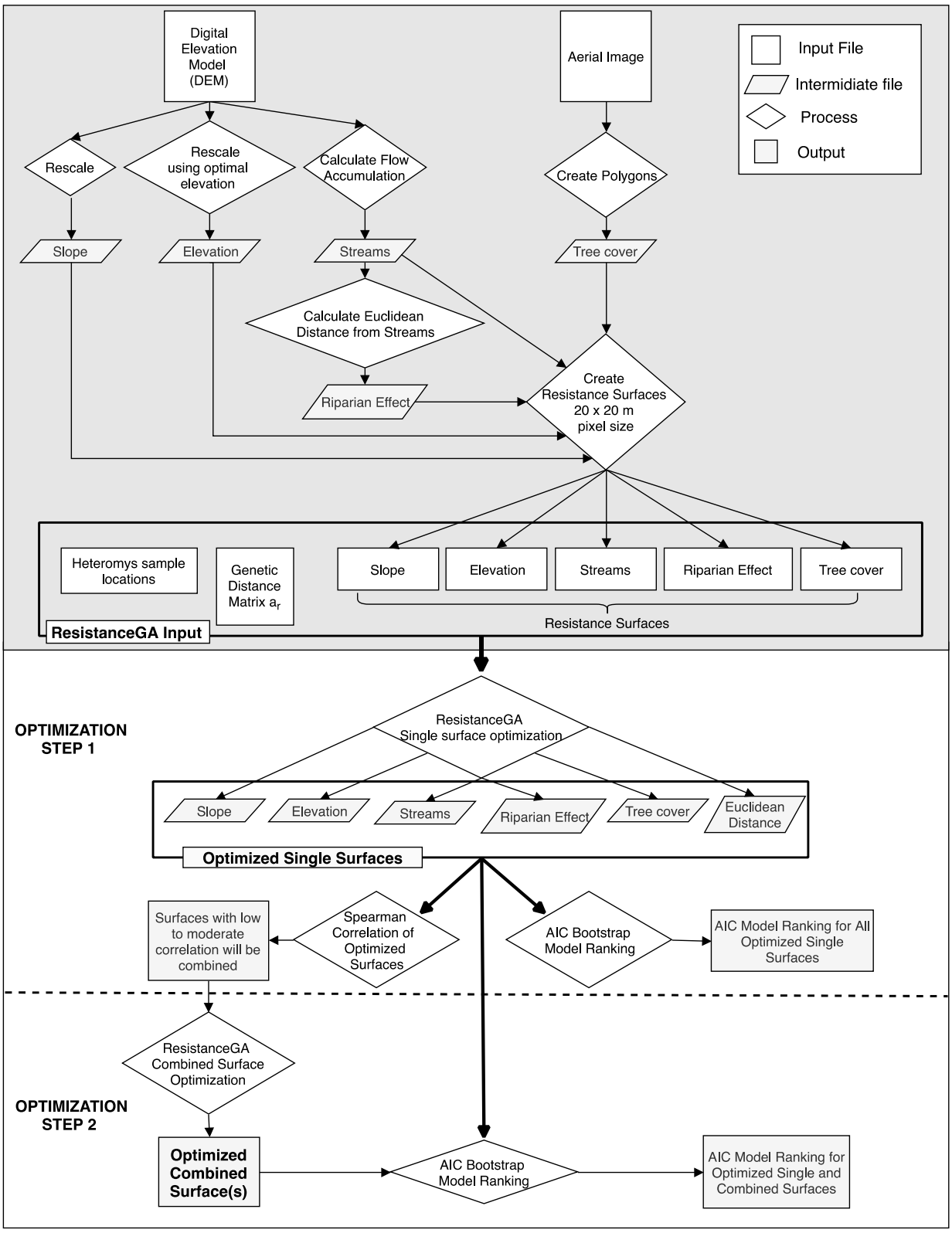


Figure 1-2 Flowchart of landscape data extraction (top, gray background) for the creation of resistance surfaces for each landscape variable and resistance surface optimization (bottom, white background).

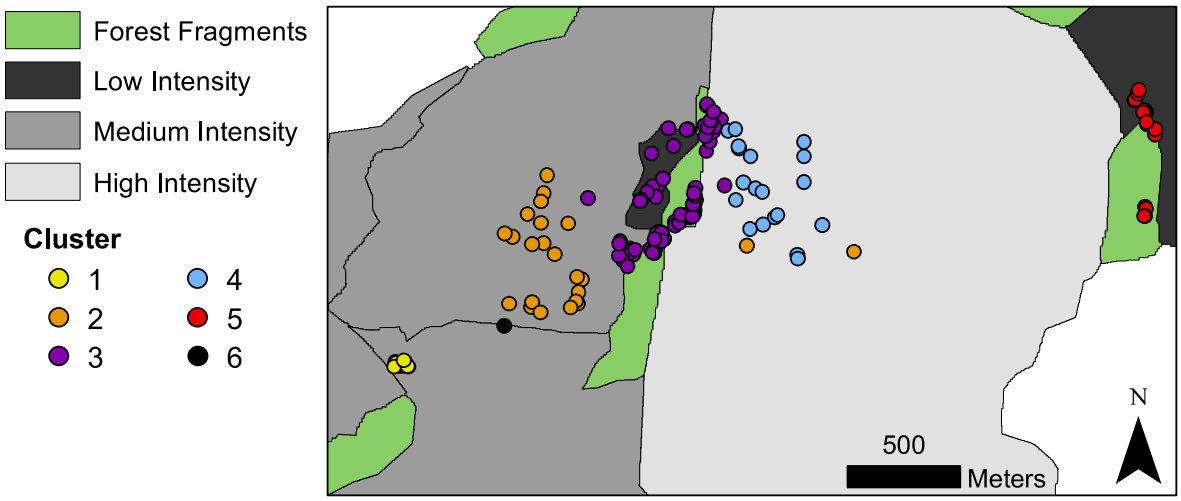


Figure 1-3 Geneland clustering results for *H. d. goldmani*. Each circle represents a sampled individual and each color a different cluster membership.

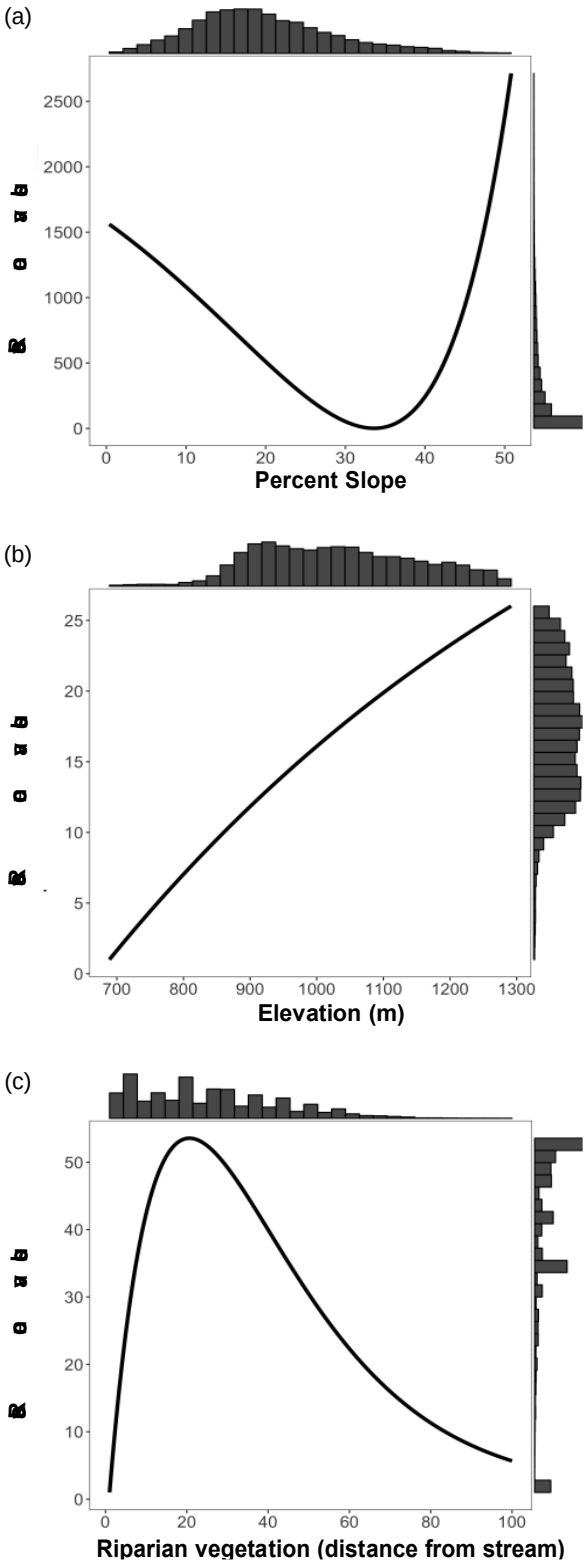


Figure 1-4 Response curves for (a) slope, (b) elevation and (c) riparian effect of single surface optimization using genetic distance for *H. d. goldmani*. Bars along the plots represent the distribution of resistance values.

References

Altieri MA (1999) Applying agroecology to enhance the productivity of peasant farming systems in Latin America. Environment, Development and Sustainability 1: 197–217

Balkenhol N, Pardini R, Cornelius C, Fernandes F, Sommer S (2013) Landscape-level comparison of genetic diversity and differentiation in a small mammal inhabiting different fragmented landscapes of the Brazilian Atlantic Forest. Conservation Genetics 14(2):355-367

Berkman LK, Nielsen CK, Roy CL, Heist EJ (2018) A landscape genetic analysis of swamp rabbits (*Sylvilagus aquaticus*) suggests forest canopy cover enhances gene flow in an agricultural matrix. Canadian Journal of Zoology 96(6):622-632

Blois JL, McGuire JL, Hadly EA (2010) Small mammal diversity loss in response to late-Pleistocene climatic change. Nature 465(7299):771-774

Bolstad P (2016). GIS fundamentals. *A first text on Geographic Information Systems*. (5^{TH} edition)

Brown JH, Heske EJ (1990) Control of a desert-grassland transition by a keystone rodent guild. Science 250(4988):1705-1707

Carver S (2010) Resistance of mammal assemblage structure to dryland salinity in a fragmented landscape. Journal of the Royal Society of Western Australia. pp. 119-128

Caudill SA, Rice RA (2016) Do Bird Friendly® Coffee Criteria Benefit Mammals? Assessment of Mammal Diversity in Chiapas, Mexico. PLOS ONE 11(11):e0165662-e0165662

Chiappero MB, Panzetta-Dutari GM, Gómez D, Castillo E, Polop JJ, Gardenal CN (2011) Contrasting genetic structure of urban and rural populations of the wild rodent *Calomys musculinus* (Cricetidae, Sigmodontinae). Mammalian Biology 76(1):41-50

Chung KP, Corlett RT (2006) Rodent diversity in a highly degraded tropical landscape: Hong Kong, South China. Biodiversity & Conservation 15(14):4521-4532

Davidson AD, Lightfoot DC (2006) Keystone rodent interactions: prairie dogs and kangaroo rats structure the biotic composition of a desertified grassland. Ecography 29(5): 755-765

DeMattia EA, Curran LM, Rathcke BJ (2004) Effects of small rodents and large mammals on Neotropical seeds. Ecology 85(8):2161-2170

Dickman CR (1999). Rodent-ecosystem relationships: a review. *Ecologically-based management of rodent pests. ACIAR Monograph*, (59), 113-133.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10: 564-567.

Fischer C, Thies C, Tscharntke T (2011) Small mammals in agricultural landscapes: Opposing responses to farming practices and landscape complexity. Biological Conservation 144(3):1130-1136

Fleming TH (1974) The Population Ecology of Two Species of Costa Rican Heteromyid Rodents. Ecology 55(3):493-510

Fleming TH (1983). Heteromys desmarestianus. *Costa Rican natural history (DH Janzen, ed.). University of Chicago Press, Chicago, Illinois*, 474-475.

Flores-Manzanero A, Luna-Bárcenas MA, Dyer RJ, Vázquez-Domínguez E (2018) Functional connectivity and home range inferred at a microgeographic landscape genetics scale in a desert-dwelling rodent. Ecology and Evolution 9(1):ece3.4762-ece3.4762

Gerlach G, Musolf K (2000) Fragmentation of Landscape as a Cause for Genetic Subdivision in Bank Voles. Conservation Biology 14(4):1066-1074

Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation--the correlated allele frequencies model revisited. Bioinformatics 24(19):2222-2228

Hanski I (1999). Metapopulation ecology. Oxford University Press.

Hooke RL, Martín-Duque JF, Pedraza J (2012) Land transformation by humans: a review.

Keeley ATH, Beier P, Gagnon JW (2016) Estimating landscape resistance from habitat suitability: effects of data source and nonlinearities. Landscape Ecology 31(9):2151-2162

Klinger R (2007) Catastrophes, disturbances and density-dependence: population dynamics of the spiny pocket mouse (Heteromys desmarestianus) in a neotropical lowland forest. Journal of Tropical Ecology 23(05):507-518

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. Conservation Genetics 7(2):295-302

Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. American Entomologist 15(3):237-240

Lomolino MV, Perault DR (2001) Island biogeography and landscape ecology of mammals inhabiting fragmented, temperate rain forests. Global Ecology and Biogeography 10(2):113-132

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. Genetics 152(4):1753-1766

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate. Trends in Ecology & Evolution 20(3):136-142

Martensen AC, Ribeiro MC, Banks-Leite C, Prado PI, Metzger JP (2012) Associations of Forest Cover, Fragment Area, and Connectivity with Neotropical Understory Bird Species Richness and Abundance. Conservation Biology 26(6):1100-1111

Martínez-Gallardo R, Sánchez-Cordero V (1993) Dietary value of fruits and seeds to spiny pocket mice, *Heteromys desmarestianus* (Heteromyidae). Journal of Mammalogy 74(2): 436-442

Moguel P, Toledo VM (1999) Biodiversity conservation in traditional coffee systems of Mexico. Conservation Biology 13(1):11-21

Nicolas V, Bryja J, Akpatou B et al (2008) Comparative phylogeography of two sibling species of forest-dwelling rodent (*Praomys rostratus* and *P. tullbergi*) in West Africa: different reactions to past forest fragmentation. Molecular Ecology 17(23):5118-5134

Olah G, Smith AL, Asner GP, Brightsmith DJ, Heinsohn RG, Peakall R (2017) Exploring dispersal barriers using landscape genetic resistance modelling in scarlet macaws of the Peruvian Amazon. Landscape Ecology 32:445-456

Otero-Jiménez B, Vandermeer JH, Tucker PK (2018) Effect of coffee agriculture management on the population structure of a forest dwelling rodent (Heteromys desmarestianus goldmani). Conservation Genetics 19(2):495-499

Oyler-McCance SJ, Fedy BC, Landguth EL (2013) Sample design effects in landscape genetics. Conservation Genetics 14(2):275-285

Passamani M, Fernandez FAS (2011) Abundance and richness of small mammals in fragmented Atlantic Forest of southeastern Brazil. Journal of Natural History 45(9-10):553-565

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19):2537-2539

Perfecto I, Vandermeer J (2002) Quality of agroecological matrix in a tropical montane landscape: Ants in coffee plantations in southern Mexico. Conservation Biology 16(1):174-182

Perfecto I, Vandermeer J (2008) Biodiversity conservation in tropical agroecosystems. Annals of the New York Academy of Sciences 1134(1):173-200

Perfecto I, Vandermeer J, Wright A (2010). Nature's matrix: linking agriculture, conservation and food sovereignty. Routledge.

Perfecto I, Vandermeer J (2010) The agroecological matrix as alternative to the land-sparing/agriculture intensification model. Proceedings of the National Academy of Sciences 107(13):5786-5791

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

Peterman WE, Connette GM, Semlitsch RD, Eggert LS (2014) Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. Molecular Ecology 23(10):2402-2413

Philpott SM, Arendt WJ, Armbrecht I, Bichier P, Diestch TV, Gordon C, Greenberg R, Perfecto I, Reynoso-Santos R, Soto-Pinto L, Tejeda-Cruz C (2008) Biodiversity loss in Latin American coffee landscapes: review of the evidence on ants, birds, and trees. Conservation Biology 22(5):1093-1105

Popescu VD, Hunter ML (2011) Clear-cutting affects habitat connectivity for a forest amphibian by decreasing permeability to juvenile movements. Ecological Applications. 21(4):1283-95.

Quintero G, Sánchez-Cordero V (1989) Estudio del área de actividad de heteromys desmarestianus (Rodentia: Heteromyidae) en una selva alta perennifolla. Anales Inst. Biol. Univ. Nac. Autón. México, Ser. Zool. pp. 28-1989

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution 49(6):1280-1283

Rogers DS, González MW (2010) Phylogenetic relationships among spiny pocket mice (*Heteromys*) inferred from mitochondrial and nuclear sequence data. Journal of Mammalogy 91(4):914-930

Rousset F (2000) Genetic differentiation between individuals. Journal of Evolutionary Biology 13(1):58-62

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol. Ecol. Resources 8: 103-106.

Russo I-RM, Sole CL, Barbato M, von Bramann U, Bruford MW (2016) Landscape determinants of fine-scale genetic structure of a small rodent in a heterogeneous landscape (Hluhluwe-iMfolozi Park, South Africa). Scientific Reports 6(1):29168-29168

Safner T, Miller MP, McRae BH, Fortin MJ, Manel S (2011) Comparison of Bayesian Clustering and Edge Detection Methods for Inferring Boundaries in Landscape Genetics. International Journal of Molecular Sciences 12(2):865-889

Santos-Filho M, Peres CA, da Silva DJ, Sanaiotti TM (2012) Habitat patch and matrix effects on small-mammal persistence in Amazonian forest fragments. Biodiversity and Conservation 21(4):1127-1147

Skellam JG (1951) Random dispersal in theoretical populations. Biometrika 38(1/2):196-218

Spear SF, Balkenhol N, Fortin MJ, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Molecular ecology 19(17):3576-3591

Vandermeer J, Carvajal R (2001) Metapopulation dynamics and the quality of the matrix. The American Naturalist 158(3):211-220

Vandermeer J, Perfecto I (2007) The agricultural matrix and a future paradigm for conservation. Conservation biology. 21(1):274-7

van Etten J (2017) *R* Package *gdistance*: Distances and Routes on Geographical Grids. Journal of Statistical Software 76(13)

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4(3):535-538

Wright S (1943) Isolation by Distance. Genetics 28(2):114-38

Wright S (1951) The genetical structure of populations. Annals of Eugenics 15(4):323-354

Zeller KA, McGarigal K, Whiteley AR (2012) Estimating landscape resistance to movement: a review. Landscape Ecology 27(6):777-797

Chapter 2 Population Connectivity Of A Generalist And A Specialist Rodent Species Across An Agricultural Intensification Gradient In A Coffee Agroecosystem

1 Introduction

Humans have been driving landscape changes for millennia. In the past 300 years technological developments have allowed humans to radically alter landscapes at a faster rate and at a much larger spatial scale than before (Ellis et al. 2013). Currently it has been estimated that 50-75% of the Earth's land surfaces have been transformed by humans (Ellis and Ramankutty 2008, Hooke et al. 2012). Species, populations and individuals are impacted by these changing environments, which in turn can affect the composition of ecological communities. In the tropics most of these land transformations have been for the expansion of agricultural production (FAO 2016). This agricultural production varies greatly in crop type, crop diversity, and management intensity, where the latter can range from high crop diversity and low external inputs (e.g., organic or traditional production) to systems with a single crop species and increased use of external inputs (i.e., pesticides, herbicides; Perfecto et al. 2010). Human driven landscape transformations (e.g., agriculture) are non-random processes that serve as a filter, favoring species, populations, and communities that can survive in modified ecosystems (Smart et al. 2006).

Ecologist have used the specialist-generalist continuum to study species responses to rapid environmental change. These categories are tightly linked to the ecological niche concept (Hutchinson 1957) and are mainly a reflection of the resource utilization capacity of individuals

(Devictor et al. 2010). A specialist species is considered to have a narrower niche, utilizing few resources and benefiting from homogeneous environments. On the other hand, a generalist species has a broader niche, utilizing many different resources and benefiting from heterogeneous environments. Based on these definitions we expect that in stable homogeneous environments specialist would thrive if they can effectively exploit the resources present. However, in changing environments generalist species are expected to persist and perform much better since they are able to utilize a range of resources. This tradeoff is known as the "jack of all trades master of none" scenario (MacArthur 1972). The specialist-generalist continuum has been a useful tool when evaluating how species responses to landscape changes vary depending on their ecological specialization.

Studies have found changes in ecological communities' local diversity and species turn-over following anthropogenic landscape changes (Olden 2006). Research has shown patterns of specialist species loss and increase of generalist species abundance as human land uses intensify (butterflies: Börschig et al. 2013, birds: Jetz et al. 2007, plants: Helm et al. 2005). Several studies have recorded functional homogenization of communities due to these changes, which can lead to the loss of ecosystem services (Clavel et al. 2010). However, most of these studies focus on species composition and density, and do not evaluate the population level responses of specialist and generalist species to changing landscapes. Studies evaluating the dispersal and population structure of species can provide valuable insight and serve as a guide for conservation management.

Technological advances and the reduction in cost of molecular analysis has allowed researchers to evaluate species population connectivity and dispersal, indirectly as a measure of

gene flow (Manel et al. 2013). This has expanded the research possibilities and made the study of non-model and elusive organisms easier.

Small mammals comprise approximately 40% of all mammal species and play many important ecological roles such as seed and fungi spore dispersers, predators, prey and grazers (Lidicker et al. 1975, Keesing 2000, Maclean et al. 2011). However, they are usually overlooked in the scientific literature, as many are small non-charismatic species (Amori and Gippolotti 2000). Some studies have found a low abundance or elimination of specialist species with increased fragmentation and agricultural intensity (Bali et al. 2007; Pardini et al. 2009). In this study we wish to add to the existing literature on tropical rodents by comparing the responses of *Heteromys desmarestianus goldmani* and *Peromyscus gymnotis* populations to agricultural management practices. *H. d. goldmani* has been classified as a forest dwelling specialist with a diet that consist mostly of seeds (Fleming 1974). *P. gymnotis* has been described as a habitat and diet generalist like most of the other species in this genus (Vázquez and Reid 2016). It is found in a range of habitat types and has a varied diet consisting of plants and insects (Vázquez and Reid 2016).

In this study we use genetic and landscape data to answer two main questions; (1) Will the patterns of population connectivity of *H. d. goldmani* (a specialist) and *P. gymnotis* (a generalist) differ? (2) Which landscape variables, if any, are promoting or impeding connectivity in these species? We expect *H. d. goldmani* to have stronger limitations to population connectivity when compared to *P. gymnotis*. We expect patterns of connectivity for *H. d. goldmani* to be mostly explained by landscape characteristics related to agricultural management intensification, while patterns for *P. gymnotis* would mostly be explained via isolation by distance.

2 Methods

2.1. Study System and Sampling

2.1.1 Study site

Coffee production in Latin America represents an ideal system for studying the effects of different management practices. In this region, coffee is produced in a variety of ways that follow an intensification gradient (Fig. 1-1). It is common to find several of these management practices represented in small areas. For this reason, we chose to do our study in the tropical montane region of Soconusco in Chiapas, Mexico (Fig. 1-1). This area is dominated by coffee production that varies in management intensity, ranging from rustic to unshaded monocultures with forest patches scattered between them (Fig. 1-1). Farms included in the study have been producing coffee for 60 to 100 years. Although management practices vary over time, these farms have had similar management practices for at least the past 20 years (Perfecto and Vandermeer 2002).

2.1.2 Study species

Small mammals have been shown to be negatively affected by human-driven landscape modifications (Umetsu and Pardini 2007, Gibson et al. 2013, Woinarski et al. 2010). Additionally, due to their short generation time, changes in their genetic structure can be perceived more rapidly. These characteristics make small mammals an ideal group in which to study the effects of anthropogenic landscape changes. Understanding their responses to agricultural management and fragmentation is also important because they are members of tropical ecosystems, serving important functional roles as seed dispersers, predators, and prey. In addition, they influence plant

structure by driving vegetation complexity and diversity through dispersal (Daily et al. 2003, Avenant 2011, Young et al 2015).

We sampled two of the most common species of small mammal in the study region. In this study we worked with *Heteromys desmarestianus goldmani*, Goldman's spiny pocket mouse, which is part of the larger *H. desmarestianus* species complex. *H. d. goldmani* are forest habitat and diet specialist with a home range of 100 m^2 (Fleming 1974). *H. desmarestianus* is the most common forest rodent in its range that extends from southern Mexico to Panama. The naked-ear deer mouse (*Peromyscus gymnotis*) is a generalist rodent that occurs in forests, grasslands, and coffee plantations in southern Mexico and Nicaragua (Vázquez and Reid 2016). Its home range is unknown, but closely related species have home ranges between 2800-4800 m^2 (Scheibe 1984). Little is known about the population structure and dispersal pattern of either species.

2.1.3 Field Sampling

H. d. goldmani and *P. gymnotis* samples were collected from 2012 to 2017. Ear tissue samples from *H. d. goldmani* and *P. gymnotis* were collected from six sites: three forest fragments and three coffee farms of various management levels that are adjacent to the forest fragments (Fig. 1-1). To facilitate continuous animal collection across this complex landscape we sampled an area of approximately 4 km x 2 km (Fig. 1-1), by following roads and trails within the farms and the forest fragments, following the methodology of Otero Jiménez et al. 2018. Ear tissue samples were preserved in 20% DMSO buffer saturated with NaCl. Mice were captured using 22.9 cm x 7.6 cm x 8.9 cm Sherman live traps. Species, sex, life stage and GPS coordinates for each individual sample were recorded. Animals were handled in accordance with the University of Michigan's Committee on Use and Care of Animals.

2.2. Genetic Data

2.2.1 SNP library preparation

Genomic DNA was extracted from ear tissue samples using the Qiagen DNeasy blood and tissue kit following the manufacture's protocol (Qiagen Inc.) and eluted in water. DNA libraries were constructed using a triple enzyme RADseq (3RAD) protocol (Glenn et al. 2017). We digested approximately 100 ng of DNA from each individual with ClaI, BamHI, and MspI restriction enzymes. We ligated all fragments to internal adapters with indexing tags of 5-8 nucleotides. A reduced cycle PCR and the standard Illumina adapter ligation protocol using Kapa LTP library preparation kits (Kapa Biosystems, Wilmington, MA) were used to construct full-length libraries. Samples were indexed using iTru5 and iTru7 primers (Glenn et al. 2016) and pooled. We used Pippin Prep (Sage Science, Beverly, MA) to size-select for fragments between 473 bp and 578 bp. We sequenced 56 *H. d goldmani* and 52 *P. gymnotis* individuals in separate (i.e., each species) 150bp pair-end Illumina HiSeq 4000 lanes at the University of Michigan DNA Sequencing Core.

2.2.2 SNP calling and filtering

First, we checked for adapter contamination and quality using FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Next, we used the *process_radtags* script from Stacks 2.1 (Catchen et al. 2013) to identify reads from each individual in the sequence data and trim all reads to 120 bp. Then we assembled *denovo* sequences for each species following the 6 step Stacks 2.1 pipeline (Catchen t al. 2013): (1) *ustacks* builds the SNP loci for each individual, (2) *cstacks* then creates a catalog of loci for all samples, (3) *sstacks* matches loci from individuals to the catalog. All these steps are done for

each read separately. (4) Next, the *tsv2bam* script was used to pull together the pair-end reads associated with each SNP and orient the data by locus. (5) The *gtsacks* program assembled and merged paired-end contigs and called variant sites in each sample. (6) Finally, we generated data sets for each species using the *population* program. These data sets contained only SNPs that were (1) found in >75% of the samples (r = 0.75); (2) had a minor allele frequency ≥ 5% (min_maf = 0.05); (3) had moderate heterozygosity ≤ 0.8; and (4) retained only a single SNP per RADtag (Catchen et al. 2013).

The generated datasets were further filtered for over clustering using VCFtools (Daneceket al. 2011). We removed sites with a total depth > 2500 and retained sites with a mean depth between 10 and 40. We verified that individuals were genotyped at ≥ 95 % of all loci. The remaining loci for each species were used for the subsequent analysis.

2.2.3 Population genomic analysis

2.2.3.1 Genetic Diversity

We calculated genome wide diversity by measuring expected heterozygosity (H_E), observed heterozygosity (H_O), nucleotide diversity (π), and inbreeding coefficient (F_{IS}) for each species separately for the filtered SNP library using the *population* function in Stacks 2.1 (Catchen et al. 2013).

2.2.3.2 Isolation by distance

To assess the influence of geographic distance on the genetic structure we calculated isolation-by-distance (IBD). We estimated pairwise Euclidean and genetic distances in Genalex

(Peakall and Smouse 2012). Next, we measured the correlation between pairwise genetic distance of all sampled individuals, and Euclidian distance using a simple Mantel test in R (R Core Team, 2018). In addition, we estimated a Mantel correlogram based on 100 m distance classes and calculated significance using Spearman correlation based on 10,000 permutations using the *ecodist* R package. These analyses were done for each species separately.

2.2.3.3 Clustering

We conducted population structure analysis using the SNPs retained after filtering across all samples. This included a Bayesian clustering analysis using both fastSTRUCTURE v1.0 (Raj et al. 2014) and TESS3 R package (Caye et al. 2016) for samples of each species separately. The admixture model of fastSTRUCTURE assumes that each individual has some proportion of ancestry originating from a number (K) of gene pools, which is reflected in the inferred ancestry coefficients (Q value). We also applied the logistic prior to the model which allows us to detect more subtle signals of genetic structure. TESS3 acts under similar assumptions but the model incorporates the geographic distance between samples, and thus it is able to detect weaker population structure. In both cases, we tested values of K ranging between 1-10, with 20 replicate runs per K. For fastSTRUCTURE, optimal K was chosen using a combination of the 'chooseK' script, and cross validation error, and for TESS3, we used cross entropy scores to determine the optimal K value. Individual ancestry coefficients were plotted using *distruct* for fastSTRUCTURE results and were plotted in R for the TESS3 results. We used geographic location information and the inferred ancestry coefficient for each individual for the best supported K to generate the population membership maps in ArcGIS. Individuals with $Q > 0.8$ for any given cluster were assigned to that cluster, all other individuals were considered

admixed. If population structure was found we calculated genetic diversity measures using the *population* function in Stacks (Catchen et al. 2013) to compare values between clusters.

2.3. Landscape Analysis

2.3.1 Landscape Data

Landscape variables – Within a 4 km long (E-W) by 2 km wide area (Fig. 1c), we tested the significance of 5 landscape features we hypothesized could influence resistance to dispersal: 1) tree cover (TC), 2) slope (S), 3) elevation (E), 4) riparian effect (RE), and 5) streams (STR; Table 2-1).

Landscape features were quantified using products derived from aerial imagery (Google 2015) and a 20 m resolution digital elevation model (DEM). Land cover was digitized manually using the heads-up digitizing method (Bolstad 2016) to map polygons of similar levels of tree cover discernable at a 1:2,000 scale. Land cover was classified into a ranking of increasing apparent habitat quality: 1) no high vegetation cover (open), 2) sun coffee plantation (no tree cover), 3) 25% tree cover coffee plantation, 4) 50% tree cover coffee plantation, 5) 75% tree cover coffee plantation, and 6) 100% tree cover (Fig. 1-2). Slope and elevation layers were developed from the DEM using appropriate geoprocessing tools (Fig. 1-2). Streams were delineated from the DEM by calculating flow accumulation and classifying any cells above a value of 100 as a stream (Fig. 1-2). We created the riparian effect layer by calculating a Euclidean distance raster from stream locations, which we used to represent lower resistance with closer proximity to a stream (Fig. 1-2). Digitization and DEM raster processing were performed with ArcGIS 10.3.1 (ESRI, 2015).

2.3.2 Resistance Surfaces

Each landscape variable layer was converted to a raster map representing the assumed resistance to movement of each variable to *H. d. goldmani* and *P. gymnotis* (Table 2-1, Fig. 1-2). These predictions of landscape effects on movement were based on the available natural history research of both study species and of other members of the *Heteromys* and *Peromyscus* genera (*Heteromys*: Fleming 1984, Martínez-Gallardo and Sánchez-Cordero 1993; *Peromyscus*: Orrock et al. 2003, Fuller et al. 2006). Landscape surfaces were rescaled to the finest possible resolution (20 m x 20 m) to reflect the hypothetical perceived scale of resistance for both rodent species.

2.4. Landscape Genetic Analysis

Most studies evaluating the effects of landscape resistance on genetic population connectivity rely on expert opinion to describe the relationship between landscape variables and species movement (Zeller et al. 2012). We used the method developed by Peterman et al. (2014) to optimize our resistance surfaces to address concerns of analysis based on expert opinion. For example, many studies using expert opinion to develop landscape resistance models treat expert opinion as empirical data, when it is not, thus making it difficult to evaluate performance (Zeller et al. 2012). In general, it is difficult to describe accurately the ecological processes being modelled in resistance analyses, and even if the processes are known, there is no guarantee that they will have significant influence on gene flow (Peterman et al. 2018). Peterman et al.'s (2014) optimization approach relies on a genetic algorithm to explore the parameter space, which seeks to maximize the relationship between pairwise landscape resistance distances (i.e., landscape variables) and pairwise genetic distances without any a priori assumptions (Peterman et al. 2018).

The optimization was done using the R package ResistanceGA (Peterman et al. 2018). Landscape surfaces were optimized in two steps for *H. d. goldmani* and *P. gymnotis* separately: (1) single resistance surface optimization, and (2) combined (i.e., multi-) surface optimization (Fig. 1-2).

For the first step we optimized each landscape surface independently, using the *commuteDistance* function exploring resistance values up to 2,500 using an 8 neighbor joining scheme to measure connectivity. The *commuteDistance* function, which calculates the commute distance or the time it takes for an individual to move from point *a* to point *b* in a particular landscape (vanEtten 2017). Our dependent variable was genetic distance and our predictor variables were the landscape resistance values. For continuous landscape variables (i.e., slope, elevation and riparian effect) we tested 9 possible transformations for the resistance relationship between both variables (i.e., linear, monomolecular, reverse monomolecular, inverse monomolecular, inverse-reverse monomolecular, Ricker, reverse Ricker, inverse Ricker, inverse-reverse Ricker). Additionally, we included pair-wise geographic distance of samples as our null model. We conducted 3 independent optimization runs for each of the landscape variables to ensure results were robust. The optimal resistance surface was identified using the Akaike Information Criterion (AIC), determined by a linear mixed effects model with a maximum likelihood population effects parameterization (MLPE; Peterman et al. 2018). We then conducted a bootstrap analysis in which 75% of the samples were randomly selected without replacement and each surface was fit to the subset of samples. For each subset of samples, the models' average rank, average weight and percentage that a surface was selected (top rank) were calculated with 10 000 iterations.

Before we continued to the second optimization step for combined surfaces, we conducted a Spearman coefficient correlation test in R between the top commute distance matrices of all optimized single surfaces that performed better than geographic distance alone. We selected variables that showed small to moderate correlation ($p \leq 0.49$), to avoid including correlated variables in the combined surface model.

For the second optimization step we performed the combined surface optimization with surfaces that showed moderate to low correlation (Fig. 1-2). Parameters for the combined surface optimization were the same as for the first optimization step. Additionally, we performed bootstrap model selection using the same parameters as for the single surface optimization in order to obtain the average rank, average model weight and the top ranked model for individual and combined surfaces. Optimization R script was adapted from Flores-Manzareno et al. (2018) to conduct this analysis.

3 Results

3.1. Sequencing and genotyping

We obtained data from 56 *H. d. goldmani* individuals and 52 *P. gymnotis*. Sequencing yielded an average of 500 million reads for each species, 9770 and 5587 SNPs after filtering *H. d. goldmani* and P. *gymnotis* respectively, with a mean coverage of 30X (range = 10-40X).

3.2. Genetic diversity

Results for genetic diversity measures show significant differences between species (Fig. 2-1, Table 2-2). *H. d. goldmani* had higher expected and observed heterozygosity when compared to *P. gymnotis* (Fig. 2-1a, b, Table 2-2). Nucleotide diversity was also higher for *H. d.*

goldmani than for *P. gymnotis* (Fig. 2-1c, Table 2-2). We found that the inbreeding coefficient (F_{IS}) did not vary significantly between species, but *P. gymnotis* has a slightly higher F_{IS} *(*Fig. 2-1d, Table 2-2).

3.3. Isolation by distance

We did not find a significant correlation between genetic and geographic distance in either of the study species when including all samples (i.e., distance classes). Mantel tests for isolation by distance for both species were not statistically significant for *H. d. goldmani* or *P.gymnotis* (Mantel r = 0.105, Mantel r = -0.128, p-value > 0.1, respectively; Table 2-2). For *H. d. goldmani* the Mantel correlogram showed a significant positive signal for IBD at distances < 300 m (Fig. 2-2a). However, *P. gymnotis* did not show significant IBD at any distance class (Fig. 2-2b)

3.4. Clustering

We analyzed population structure using two different Bayesian methods, fastSTRUCTURE (Raj et al. 2014) and TESS3 (Cave et al. 2012). Results for both methods yielded similar patterns of structure for both species. For *H. d. goldmani* fastSTRUCTURE results supported the presence of 1 to 3 clusters (K), while TESS3 showed support for K of 1 or 2, with K = 2 having a slightly lower cross-entropy score than K = 1, representing a higher likelihood (Fig. 2-3). We evaluated the plotted individual cluster membership for each of the potential best supported K values for each method (Fig. 2-4a). For fastSTRUCTURE the K = 2 included a ghost cluster, and all individuals were assigned to a single cluster, and so we did not include the plot. The results show an overlapping pattern across both methods for the assignment

of 2 clusters (Fig. 2-4a, b). The first cluster is composed of all individuals in the high, medium and low intensity coffee farms, and individuals from the middle forest fragment (Fig. 2-4c). The second cluster is composed of individuals sampled in the eastern forest fragment at the border of the high intensity coffee farm. To further explore the differentiation between clusters we compared the genetic diversity measures for each cluster and found no significant differences (Fig. 2-4d). Only nucleotide diversity was significantly different between clusters, but this difference is small (Fig. 2-4d). These results suggest subtle population structure in *H. d. goldmani*.

P. gymnotis fastSTRUCTURE analysis showed support for K = 1 - 4, while the TESS results identified K = 1 as the optimal K (Fig. 2-5a). When plotting the individual clusters assignments, we also found the presence of ghost clusters at all K values for fastSTRUCTURE (Fig. 2-5b). Assignment of individuals for K = 3 and K = 4 showed no spatial pattern. Because the two clustering methods do not show similar clustering patterns, results suggest that there is no detectable genetic structure within our sample of *Peromyscus*. Individuals assigned to different clusters in fastSTRUCTURE could indicate recent dispersal events.

3.5. Landscape genetics

Model selection results for *H. d. goldmani* and *P. gymnotis* show that slope (S) and elevation (E) resistance surfaces were the best-supported models based on 10,000 bootstrap replicates (Table 2-3). Elevation was the best supported surface for *H. d. goldmani* (top model 75% of the time; Table 2-3) and slope was the best supported surface for *P. gymnotis* (77% of the time; Table 2-3).

Results show that for *H. d. goldmani* all optimized surfaces, except for streams (STR), ranked higher than the geographic distance surface (Table 2-3). This finding supports the non-significant IBD results for *H. d. goldmani* showing that geographic distance does not have a strong influence in maintaining population connectivity. Elevation (75%) and slope (19%) were the two best supported surfaces, followed by tree cover (1.98%) and riparian effect (1.68%; Table 2-3). We found that for *P. gymnotis* only slope (77%) and elevation (19%) ranked higher than the distance surface (2.3%; Table 2-3). This suggests that the measured landscape variables related to agricultural production (e.g., tree cover and riparian effect), as well as geographic distance, have little influence on *Peromyscus* population connectivity.

The resistance patterns for each optimized surface varied between species (Fig. 2-6). The optimized slope surface for *H. d. goldmani* showed a range of low resistance values between 30-40% slope (Fig. 2-6a), while the pattern for *P. gymnotis* was the opposite with 30-40% slope representing the highest resistance in the optimized surface (Fig. 2-6b). For the elevation optimized surface *H. d. goldmani* showed a pattern of lower resistance values at elevations of 700-800 m. Additionally, the highest assigned resistance values for elevation was of approximately 200 at elevations outside of 800 m (Fig. 2-6c), which are lower than the highest assigned values for the Peromyscus elevation surface (Fig. 2-6d). The optimized elevation surface for *P. gymnotis* had low resistance values at a broader elevation range, from 900-1100 m, and much higher resistance values outside of 800 m for *Heteromys* (lower elevation: 400 resistance value, higher elevation: 800 resistance value; Fig. 2-6d). These results demonstrate how the same variables are important for the population connectivity of both species, but the direction of this response differs.

Tree cover and riparian effect were the two landscape surfaces tested that are directly influenced by management practices. Low intensity coffee management practices allow for higher tree cover in the farms and less management and production around water bodies, like streams and ravines. For the generalist species, *P. gymnotis*, tree cover was the lowest ranked surface when trying to explain genetic structure (Table 2-3) suggesting that *Peromyscus* population connectivity can be maintained at varying levels of cover. On the other hand, for *H. d. goldmani*, the specialist species, tree cover ranked as the third top model, suggesting that tree cover could be influencing connectivity. The resulting optimized tree cover surface for each species showed opposite patterns (Fig. 2-6). *Heteromys* shows tree cover > 50% to have the lowest resistance to connectivity with values any lower than 50% showing high resistance (Fig. 2-6). For *Peromyscus*, the resistance values have less variation showing areas of 0-100% tree cover having relatively low resistance values (Fig. 2-6).

We tested for correlations between all optimized landscape surfaces using Spearman's tau for each species. For *Heteromys* all surfaces showed some degree of correlation between each other (Table 2-4). Only one surface pair (i.e., slope and riparian effect) of the 15 pairs showed moderate correlation (0.3 Spearman tau), all other surface pairs showed strong correlations (> 0.5 Spearman tau; Table 2-4). In the subsequent optimization step we created only a single combined surface with slope and riparian effect, the two variables that showed lower degree of correlation. For *Peromyscus* we found more variation in degree of correlation between pairs (Table 2-4). The two variable pairs that showed no correlation included: (1) streams and tree cover, and (2) riparian effect and streams. Because all variables except slope and elevation ranked lower than geographic distance for *Peromyscus* we only modeled a single combined surface using elevation and slope.

Results from the combined surface optimization for *Heteromys* showed the combined surface to be the third best supported model at 10% (Table 2-5), after slope (15%) and elevation (72%; Table 2-5). For *Peromyscus* the combined surface also ranked third, being the top model 0.5% of the time (Table 2-5) while slope (78%) and elevation (17%) continued to be the top two models.

4 Discussion

As part of the ecosystem humans have and will continue to modify the Earth's environment. Agricultural production, both its expansion and intensification, is one of the most significant drivers of anthropogenic change. Agriculture's effect on ecosystems can range from maintaining most of the ecological community structure, to homogenization and species turn-over. It is important to study species responses to different production schemes in order to identify management and production practices that will maintain productive agricultural landscapes that sustain biodiversity and maintain ecosystem services. Additionally, we know that species level of specialization in resources and habitat requirements affect their responses to environmental changes. This study explored the population connectivity of two rodent species, with varying levels of ecological specialization, within a coffee agroecosystem composed of farms varying in management practices.

Using genetic and landscape data we sought to answer these two questions: (1) Do patterns of population connectivity of *H. d. goldmani* (a specialist) and *P. gymnotis* (a generalist) differ? (2) Which landscape variables, if any, are promoting or impeding connectivity in these species? We found different patterns of population connectivity and responses to landscape changes in each species.

Patterns of population connectivity

The presence of genetic structure in *H. d. goldmani*, suggests that there are limitations to population connectivity. Geographic distance had a weak but significant correlation with genetic distance but at small distance classes (< 300 m). The two genetic clusters identified were separated by high intensity coffee productions (i.e., sun coffee). These results suggest that this coffee production scheme may serve as a barrier to movement and decrease population connectivity of *H. d. goldmani.* In contrast, we found *P. gymnotis* individuals-maintained population connectivity supported by the absence of population structure based on the genetic data, and no significant role of geographic distance. These results differed from our expectations, since we predicted IBD to be the main contributor to the generalist species population structure. Studies have found strong patterns of isolation by distance in temperate regions for *Peromyscus* species (Moscarella et al. 2019, Howell et al. 2017). Studies on *P. leucopus* revealed IBD across the great lakes region (Moscarella et al. 2019), and in agricultural landscapes at distances > 1800 m (Howell et al. 2017). However, these studies were conducted at larger spatial scales and it is possible that our study area was not large enough to capture the influence of IBD on *P. gymnotis*.

For forest specialist species, studies have found similar patterns of IBD as those reported here (Otero Jiménez et al. 2018, Cobo-Simon et al. 2018). Cobo-Simon et al. (2018) reported that specialist species (*H. desmarestianus* and *P. mexicanus*) showed weak and non-significant patterns of IBD, but significant genetic or population structure in a tropical forest reserve in Central America. The influence of isolation by distance is overpowered by the effects of the landscape on population structure.

Genetic diversity also differed between species. *H. d. goldmani* showed higher levels of heterozygosity and nucleotide diversity when compared to *P. gymnotis*. Small isolated populations are expected to have low levels of genetic diversity and high levels of inbreeding (Melbourne and Hastings 2008) which can threaten long-term population survival. If coffee production was having this effect on *Heteromys* we would have expected lower genetic diversity than what we observed. One possibility to explain this pattern is the presence of *Heteromys* genetic structure observed. However, this is unlikely since both genetic clusters were found to have similar levels of genetic diversity. These findings suggest that despite the genetic structure found, there is strong gene flow between *Heteromys* individuals in the coffee agroecosystem. The lower values of genetic diversity for *Peromyscus* did not support our predictions, since we expected high gene flow and population connectivity, and high genetic diversity. These patterns could be driven by many different factors, such as competitive exclusion or a recent colonization event.

Landscape influences on population connectivity

Our results showed that, out of the 5 landscape characteristics evaluated, elevation and slope were the most correlated with the patterns of population connectivity in both species. However, the relationship between each species and the variables were markedly different, highlighting the ecological differences between these species. Our results show that *H. d. goldmani* connectivity is higher at a lower elevation and over a narrower range than that of *P. gymnotis* despite their overlapping elevation range (*H. desmarestianus* 200-2400 m; *P. gymnotis* 0-1700 m Reid 2009). This pattern could be explained by the fact that the sampled high intensity farm is located at the highest elevation in the study area (900-1100 m). *Peromyscus* is able to

thrive in this environment while *H. d. goldmani* is restricted to areas with lower management intensification that are coincidentally located at lower elevations. Another interesting pattern found is that *P. gymnotis* population connectivity is correlated with lower percent slope (20%), when compared to that of *H. d. goldmani* (30-40%). We believe that this could also be driven by coffee production practices, where flatter areas tend to be more densely planted and managed than steeper areas. For this reason, *Heteromys* might move through steeper landscapes and avoid intense production areas with low cover. This is supported by our findings where areas with higher tree cover promote connectivity of *H. d. goldmani*. These results can suggest the presence of competitive exclusion of *Peromyscus* from the forested areas in the coffee agroecosystem.

Studies evaluating the effect of agricultural practices on small mammal populations show varying responses (Pardini et al. 2004, Arce-Peña et al. 2019, White et al. 2012). Most studies show a general pattern of agricultural intensification increasing the density and presence of a few generalist species (White et al. 2012, Silva et al. 2005), while specialist species decrease with surrounding landscape intensification and reduced forest (Arce-Peña et al 2019, Silva et al. 2005), lowering species richness in high intensity agricultural areas. These patterns support our findings of higher population connectivity for *P. gymnotis*, the generalist species, and its lack of response to the agricultural landscape features measured in this study. Few studies exist that evaluate the effect of agricultural landscape on rodent population connectivity. One of these studies conducted in a temperate region found that larger landscape features, such as roads and rivers, are barriers to connectivity for *Peromyscus leucopus*, a generalist species (Howell et al. 2017). The effects of agriculture on the population connectivity of specialist rodents has mostly been studied in a binary habitat-non habitat context. These studies have evaluated the influences of forest fragment size, edges and distance to other fragments on connectivity, and show that

smaller, isolated fragments increase population differentiation (Balkenhol et al. 2013, Arce-Peña et al. 2019). *H. d. goldmani* population structure in our study supports the patterns found in these studies, since the agricultural matrix is reducing connectivity.

This study highlights the impact agricultural management practices can have on population connectivity. However, one of the main limitations to this type of study, specifically in tropical regions, is the availability of landscape data and the spatial grain at which it is collected. Future studies would benefit from collecting microclimatic data and other measures that can create a better profile of the low, medium and high intensity farms at a level that is relevant to rodents, as done by Flores-Manzanero et al (2018). This information could provide greater insights into the relationships between and among elevation, slope, and management, and their relative effect on the patterns observed. We also were limited by a small sample size in this study. Increasing the sampling size might reveal other patterns that we were not able to detect. We hope that as technology progresses and sequencing becomes more accessible this will no longer be a limitation in the near future.

It has been well established that human dominated environments reduce animal movement and connectivity (Tucker et al. 2018). In this study we explored rodent population connectivity within a coffee production landscape composed of farms with varying management practices. The study species, *H. d. goldmani* and *P. gymnotis* are able to survive within the coffee agricultural matrix. *Peromyscus*, the generalist species, maintains connectivity in this landscape while, *H. d. goldmani*, a forest specialist, shows signs of reduced connectivity that appears to be driven by increase in management intensification (e.g., reduced tree cover). However, genetic diversity patterns suggest that *H. d. goldmani* is able to migrate through the landscape

maintaining high levels of genetic diversity, while P. gymnotis has lower levels of genetic diversity despite the absence of genetic structure. More research is needed to understand the drivers of these patterns. These differences in response highlight the importance of sampling species representative of varying degrees of ecological specialization when studying the effects of environmental change. This is especially important for sustainability and conservation management planning. We hope that our study serves as an example of the potential that genetic and landscape data have to inform agricultural and conservation management practices. Small mammals are key components of all ecosystems and can be used as a measure of ecosystem health. This study provides novel information about tropical rodent populations that can potentially be used in future studies to measure the success of agricultural and conservation practices in maintaining connectivity.

5 Tables

Table 2-1 Expected *H. d. goldmani* and *P. gymnotis* resistance to movement for each landscape variable. We expect *Heteromys* to show stronger responses to all variables when compared to *Peromyscus*

Landscape Variable	Variable Type	Expected Resistance
Euclidean Distance		Isolation by distance. Greater resistance when further apart
Elevation	Continuous	Greater resistance as you move away from the optimal elevation
Slope	Continuous	Greater resistance as the percent slope increases
Riparian Effect	Continuous	Greater resistance further from stream edge
Tree Cover	Categorical	Greater resistance with lower tree cover
Streams	Categorical	Streams as barriers to movement

Table 2-2 Number of samples, number of SNP markers, genetic diversity measures and IBD correlation for *H. d. goldmani* and *P. gymnotis*. Values in parenthesis are the standard error.

	Species	
	Heteromys	*Peromyscus*
Samples	56	52
SNPs	9770	5587
Observed Heterozygosity	0.24712 (± 0.0014)	0.11159 (± 0.0016)
Expected Heterozygosity	0.27534 (± 0.0013)	0.14159 (± 0.0019)
Inbreeding Coefficient	0.1092 (± 0.053)	0.15562 (± 0.054)
Nucleotide diversity	0.27815 (± 0.0013)	0.14325 (± 0.0019)
Isolation-by-Distance	0.105 ($p > 0.1$)	-0.128 ($p > 0.1$)

Table 2-3 Model selection results for the generalized linear mixed -effects models optimized on genetic distance for (top) *H. d. goldmani* and (bottom) *P. gymnotis*. K is the number of parameters for each model.

Surface	K	Equation	AIC	Average Weight	Average Rank	Top Model %
Elevation	4	Inverse Ricker	9604.4	0.687	1.40	75.19
Slope	4	Inverse-Reverse Ricker	9611.2	0.195	2.62	19.77
Tree Cover	7	NA	9612.8	0.047	3.42	1.97
Riparian Effect	4	Monomolecular	9616.9	0.027	4.51	1.68
Distance	2	NA	9617.8	0.021	5.01	0.77
Streams	3	NA	9615.9	0.023	4.05	0.62

Surface	K	Equation	AIC	Average Weight	Average Rank	Top Model %
Slope	4	Reverse Ricker	6369.144	0.72520118	1.6714	77.57
Elevation	4	Inverse-Reverse Ricker	6375.327	0.10132819	2.8754	18.77
Distance	2	NA	6375.829	0.0541534	3.2369	2.49
Streams	3	NA	6375.913	0.04662508	3.7551	0.33
Riparian Effect	4	Inverse Ricker	6376.015	0.04637222	3.9836	0.71
Tree Cover	7	NA	6377.277	0.02631992	5.4776	0.13

Table 2-4 Spearman correlation results for all pairs of resistance surfaces and strength of correlation for (top) *H. d goldmani* and (bottom) *P. gymnotis*. Slope (S), tree cover (TC), elevation (E), riparian effect (RE), and streams (STR).

Surfaces	Spearman rho	P	Correlation
S + RE	0.338	< 2.2e-16	Moderate
S + E	0.515	< 2.2e-16	High
S + STR	0.536	< 2.2e-16	High
S + D	0.565	< 2.2e-16	High
S + TC	0.586	< 2.2e-16	High
RE + E	0.605	< 2.2e-16	High
TC + RE	0.660	< 2.2e-16	High
E + STR	0.675	< 2.2e-16	High
E + D	0.708	< 2.2e-16	High
TC + STR	0.769	< 2.2e-16	High
TC + D	0.815	< 2.2e-16	High
RE + D	0.865	< 2.2e-16	High
TC + E	0.866	< 2.2e-16	High
STR + RE	0.917	< 2.2e-16	High
STR + D	0.964	< 2.2e-16	High

Surfaces	Spearman rho	P	Correlation
STR + RE	-0.0235	0.24	No Correlation
TC + STR	0.0172	0.39	No Correlation
TC + E	0.0424	0.04	Low
S + E	0.0903	7.62E-06	Low
E + RE	0.2178	< 2.2e-16	Moderate
S + RE	0.2344	< 2.2e-16	Moderate
TC + RE	0.3306	< 2.2e-16	Moderate
E + STR	0.3330	< 2.2e-16	Moderate
TC + D	0.3355	< 2.2e-16	Moderate
RE + D	0.4380	< 2.2e-16	High
S + TC	0.5126	< 2.2e-16	High
E + D	0.5275	< 2.2e-16	High
S + D	0.5608	< 2.2e-16	High
S + STR	0.5707	< 2.2e-16	High
STR + D	0.6581	< 2.2e-16	High

Table 2-5 . Model selection results for both individual and composite surfaces for (top) *H. d. goldmani* and (bottom) *P. gymnotis*

Surface	K	Equation	AIC	Average Weight	Average Rank	Top Model %	Surfaces
Elevation	4	Inverse Ricker	9604.4	0.628	1.64	71.65	Elevation
Slope	4	Inverse-Reverse Ricker	9611.1	0.145	3.28	15.07	Slope
Combined	7	NA	9609.8	0.132	2.54	9.06	Slope and Riparian Effect
Tree Cover	7	NA	9612.8	0.037	4.15	1.59	Tree Cover
Riparian Effect	4	Monomolecular	9616.8	0.022	5.46	1.38	Riparian Effect
Distance	2	NA	9617.7	0.017	5.96	0.67	Distance
Streams	3	NA	9615.9	0.019	4.98	0.58	Streams

Surface	K	Equation	AIC	Average Weight	Average Rank	Top Model %	Surfaces
Slope	4	Reverse Ricker	6369.2	0.7132	1.67	78.08	Slope
Elevation	4	Inverse-Reverse Ricker	6375.5	0.0902	3.46	17.41	Elevation
Combined	7	NA	6375.2	0.0388	3.49	0.48	Slope and Elevation
Distance	2	NA	6376.0	0.0492	3.91	2.78	Distance
Streams	3	NA	6376.1	0.0424	4.43	0.25	Streams
Riparian Effect	4	Inverse Ricker	6376.2	0.0420	4.70	0.82	Riparian Effect
Tree Cover	7	NA	6377.4	0.0241	6.33	0.18	Tree Cover

6 Figures

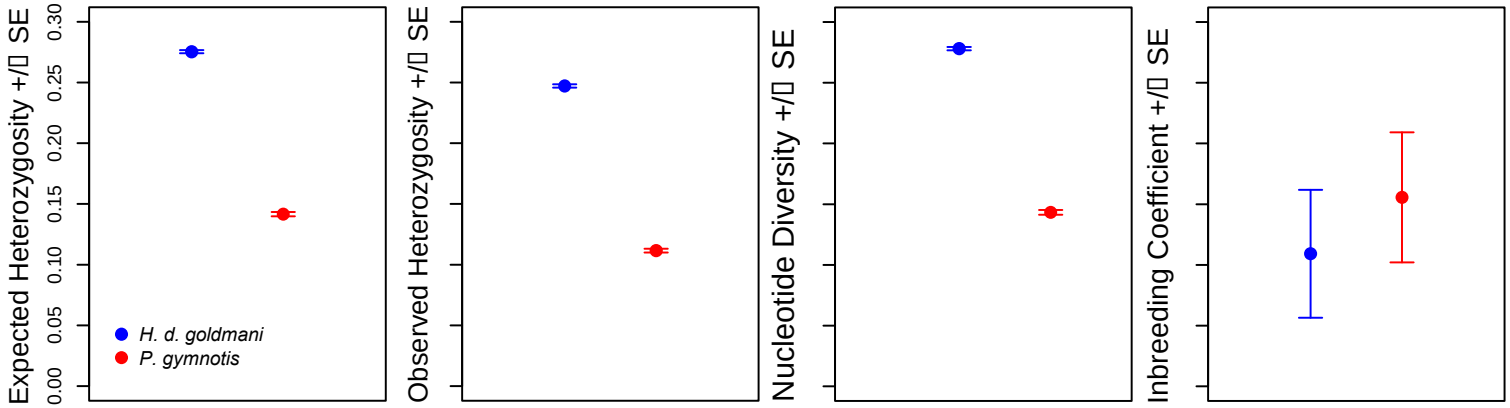


Figure 2-1 Genetic diversity measures for *H. d. goldmani* (blue) and *P. gymnotis* (red). Mean and standard error for each measure are included.

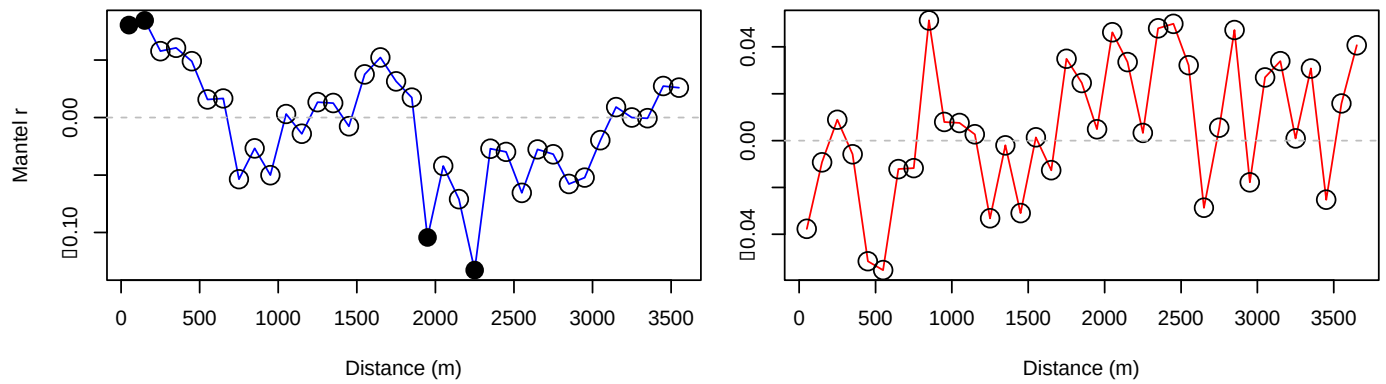


Figure 2-2 Mantel correlogram for geographic (Euclidian) and genetic distances calculated at 100 m distance classes for *H. d. goldmani* (blue) and *P. gymnotis* (red). Significant values ($p < 0.05$) are shown in black.

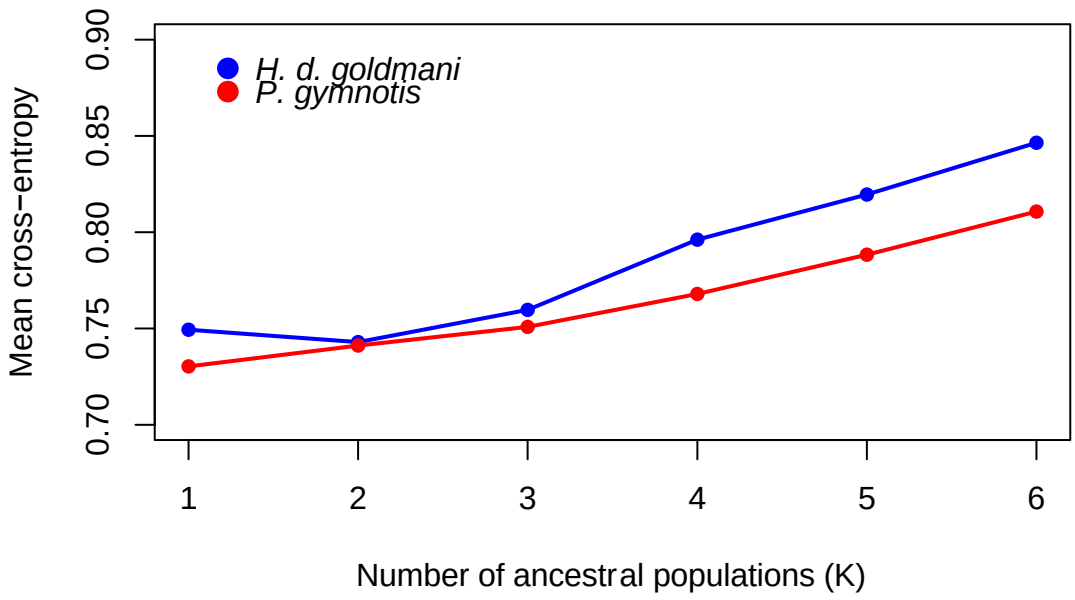


Figure 2-3 Mean cross-entropy measures across all 10 replicates for each K value for *Heteromys* and *Peromysc*

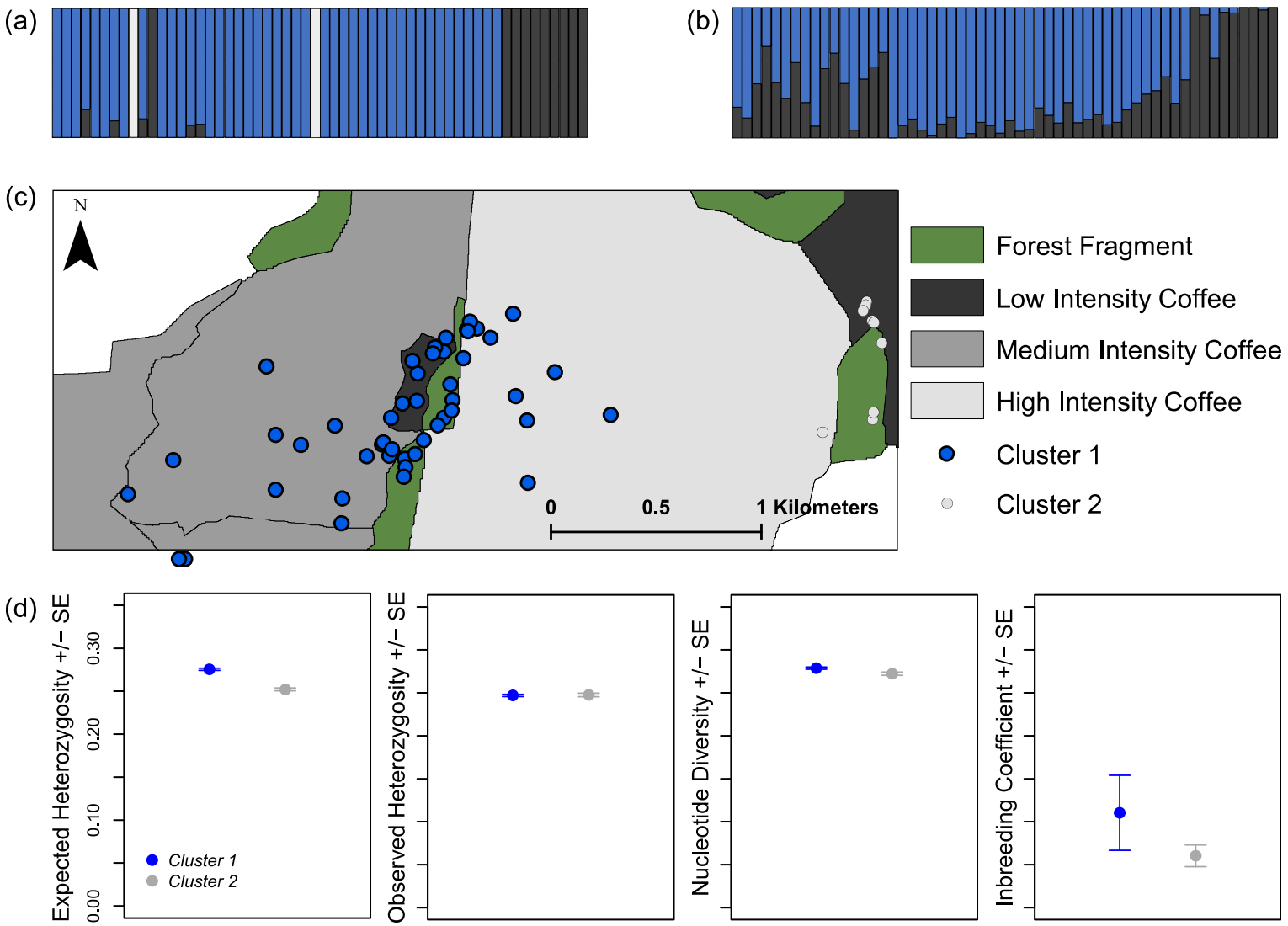


Figure 2-4 Clustering results for *H. d. goldmani* individuals. (a) fastSTRUCTURE individual assignments for K=3, and (b) for TESS K=2. (c) Map of individual cluster assignments for *H. d. goldmani*. (d) Genetic diversity measures for *H. d. goldmani* clusters. For individual cluster assignment, each bar represents an individual and colors represent the proportion membership to each cluster.

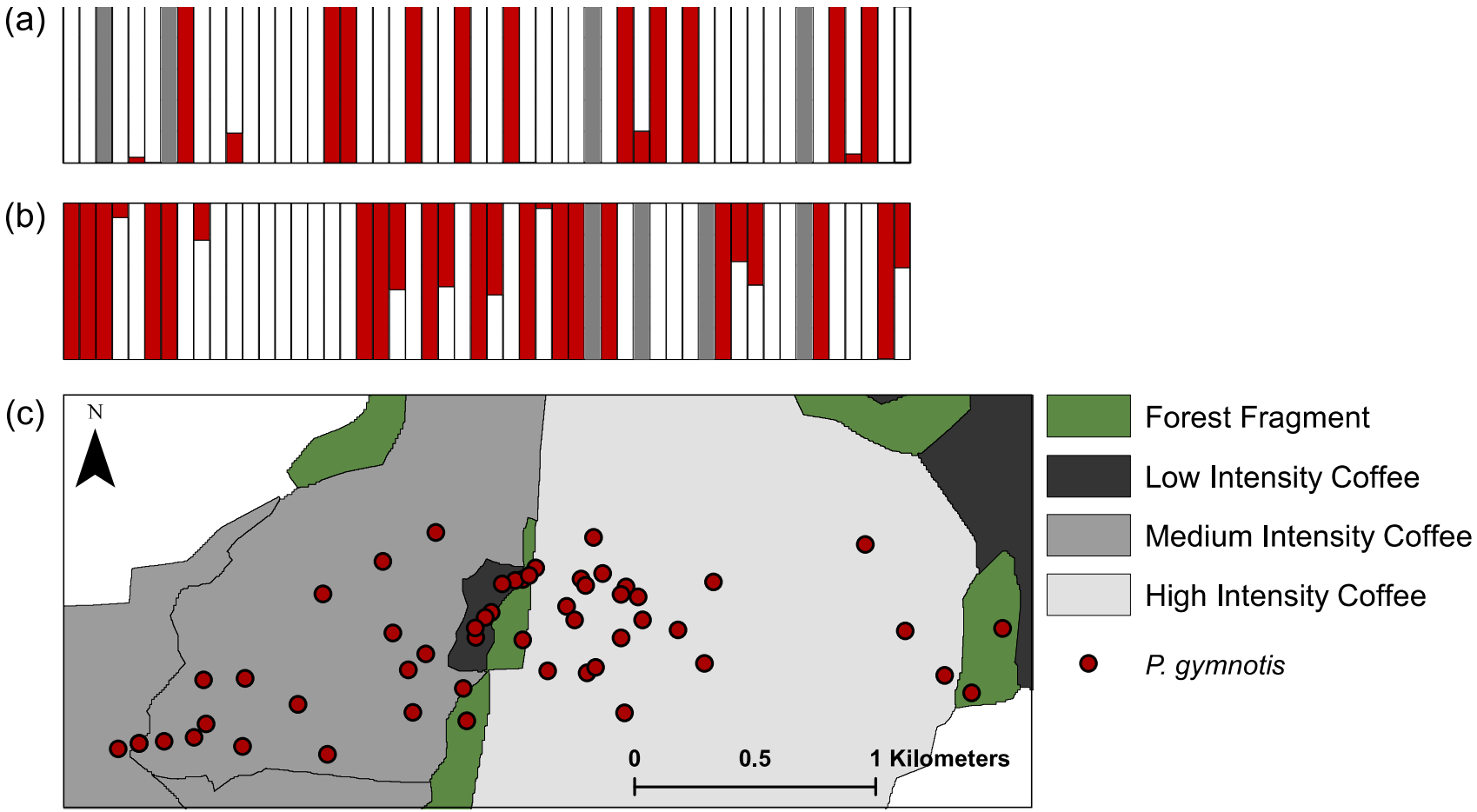


Figure 2-5 *P. gymnotis* individual assignments for (a) fastSTRUCTURE K=3, and (b) K = 4. (c) Map of *P. gymnotis* individuals. For individual cluster assignment, each bar represents an individual and colors represent the proportion membership to each cluster

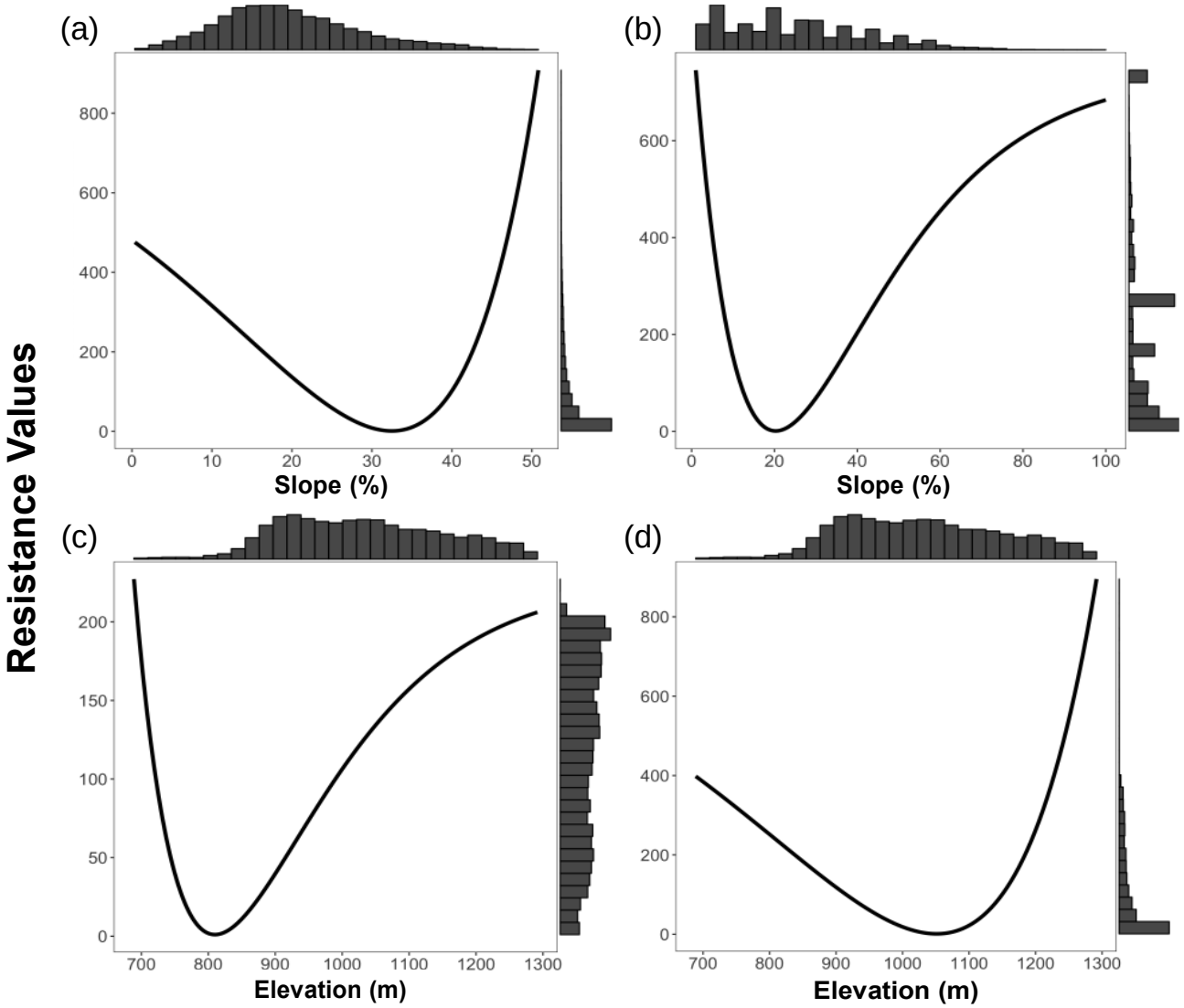


Figure 2-6 *Heteromys* (left) and *Peromyscus* (right) response curves for (a, b) elevation, (c, d) slope and (e, f) riparian effect of single surface optimization using genetic distance for each species. Bars along the plots represent the distribution of resistance values

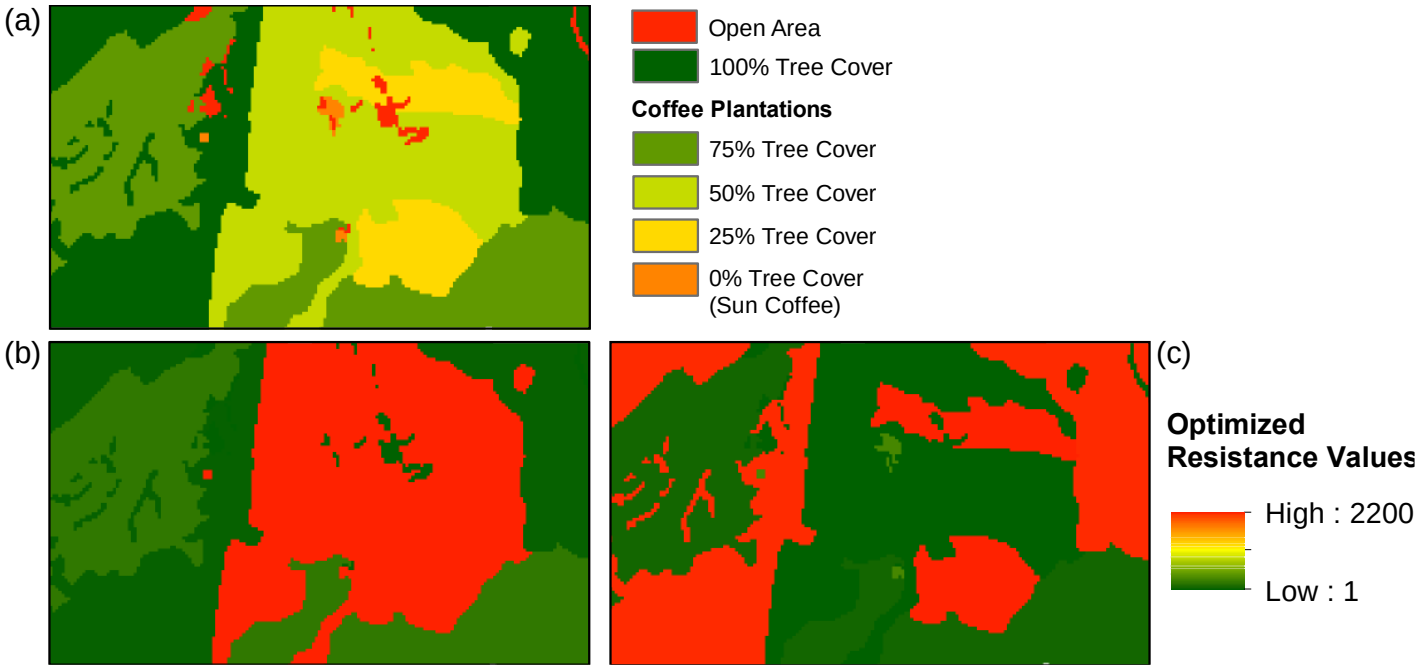


Figure 2-7 Raster map surfaces for tree cover: a) original input surface for ResistanceGA, b) resulting ResistanceGA surface for the best model of tree cover based on genetic distance for *H. d. goldmani*, and c) best model of tree cover for *P. gymnotis*

References

Amori, G., & Gippoliti, S. (2000). *What do mammalogists want to save? Ten years of mammalian conservation biology*. *Biodiversity and Conservation* (Vol. 9). Retrieved from https://link-springer-com.proxy.lib.umich.edu/content/pdf/10.1023%2FA%3A1008971823774.pdf

Arce-Peña, N. P., Arroyo-Rodríguez, V., San-José, M., Jiménez-González, D., Franch-Pardo, I., Andresen, E., & Ávila-Cabadilla, L. D. (2019). Landscape predictors of rodent dynamics in fragmented rainforests. *Biodiversity and Conservation*, *28*(3), 655–669. https://doi.org/10.1007/s10531-018-1682-z

Avenant, N. (2011). The potential utility of rodents and other small mammals as indicators of ecosystem "integrity" of South African grasslands. *Wildlife Research*, *38*(7), 626–639. https://doi.org/10.1071/wr10223

Balkenhol, N., Pardini, R., Cornelius, C., Fernandes, F., & Sommer, S. (2013). Landscape-level comparison of genetic diversity and differentiation in a small mammal inhabiting different fragmented landscapes of the Brazilian Atlantic Forest. *Conservation Genetics*, *14*(2), 355–367. https://doi.org/10.1007/s10592-013-0454-2

Bali, A., Kumar, A., & Krishnaswamy, J. (2007). The mammalian communities in coffee plantations around a protected area in the Western Ghats, India. *Biological Conservation*, *139*(1–2), 93–102. https://doi.org/10.1016/j.biocon.2007.06.017

Bolstad, P. (2016). GIS fundamentals. *A first text on Geographic Information Systems*. (5TH edition)

Börschig, C., Klein, A. M., von Wehrden, H., & Krauss, J. (2013). Traits of butterfly communities change from specialist to generalist characteristics with increasing land-use intensity. *Basic and Applied Ecology, 14*(7), 547-554.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Mol Ecol*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. (2016). TESS3: fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, *16*(2), 540–548. https://doi.org/10.1111/1755-0998.12471

Cobo-Simón, I., Méndez-Cea, B., Portillo, H., Elvir, F., Vega, H., Gallego, F. J., & Fontecha, G. (2018). Testing the effectiveness of conservation management within Biosphere Reserves: the case of Mexican Deer Mouse (*Peromyscus mexicanus*) as a bioindicator. *Integrative Zoology*. https://doi.org/10.1111/1749-4877.12371

Daily, G. C., Ceballos, G., Pacheco, J., Suzan, G., & Sanchez-Azofeifa, A. (2003). Countryside biogeography of neotropical mammals: Conservation opportunities in agricultural landscapes of Costa Rica. *Conservation Biology*, *17*(6), 1814–1826. https://doi.org/10.1111/j.1523-1739.2003.00298.x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., … Mouquet, N. (2010). Defining and measuring ecological specialization. *Journal of Applied Ecology*, *47*(1), 15–25. https://doi.org/10.1111/j.1365-2664.2009.01744.x

Ellis, E. C., & Ramankutty, N. (n.d.). *Anthropogenic biomes of the world*. Retrieved from https://esajournals-onlinelibrary-wiley-com.proxy.lib.umich.edu/doi/pdf/10.1890/070062

Ellis, E. C., Kaplan, J. O., Fuller, D. Q., Vavrus, S., Klein Goldewijk, K., & Verburg, P. H. (2013). Used planet: A global history. *Proceedings of the National Academy of Sciences*, *110*(20), 7978–7985. https://doi.org/10.1073/pnas.1217241110

FAO (2016) State of the World's Forests 2016. Forests and agriculture: land-use challenges and opportuni- ties. FAO, Rome

Fleming, T.H. (1974) The Population Ecology of Two Species of Costa Rican Heteromyid Rodents. Ecology 55(3):493-510

Fleming, T.H. (1983). Heteromys desmarestianus. *Costa Rican natural history (DH Janzen, ed.). University of Chicago Press, Chicago, Illinois*, 474-475.

Flores-Manzanero, A., Luna-Bárcenas, M. A., Dyer, R. J., & Vázquez-Domínguez, E. (2018). Functional connectivity and home range inferred at a microgeographic landscape genetics scale in a desert-dwelling rodent. *Ecology and Evolution*, *9*(1), ece3.4762. https://doi.org/10.1002/ece3.4762

Fuller, T., Munguía, M., Mayfield, M., Sánchez-Cordero, V., & Sarkar, S. (2006). Incorporating connectivity into conservation planning: A multi-criteria case study from central Mexico. *Biological Conservation*, *133*(2), 131–142. https://doi.org/10.1016/J.BIOCON.2006.04.040

Gibson, L., Lynam, A. J., Bradshaw, C. J. A., He, F. L., Bickford, D. P., Woodruff, D. S., … Laurance, W. F. (2013). Near-Complete Extinction of Native Small Mammal Fauna 25 Years After Forest Fragmentation. *Science*, *341*(6153), 1508–1510. https://doi.org/10.1126/science.1240495

Glenn, T. C., Nilsen, R. A., Kieran, T. J., Finger, J. W., Pierson, T. W., Bentley, K. E., … Faircloth, B. C. (2016). Adapterama I: Universal Stubs and Primers for Thousands of Dual-Indexed Illumina Libraries (iTru & iNext). *BioRxiv*, 049114. https://doi.org/10.1101/049114

Glenn, T., Bayona-Vásquez, N., Kieran, T., Pierson, T., Hoffberg, S., Scott, P., … Faircloth, B. (2017). Adapterama III: Quadruple-indexed, triple-enzyme RADseq libraries for about $1USD per Sample (3RAD). *BioRxiv*, 205799. https://doi.org/10.1101/205799

Helm, A., Hanski, I., & Partel, M. (2005). Slow response of plant species richness to habitat loss and fragmentation. *Ecology Letters*, *0*(0), 051109031307003. https://doi.org/10.1111/j.1461-0248.2005.00841.x

Hooke, R. L., & Martín-Duque, J. F. (2012). Land transformation by humans: A review. *GSA Today*, *12*(12), 4–10. https://doi.org/10.1130/GSAT151A.1

Howell, P. E., Koen, E. L., Williams, B. W., Roloff, G. J., & Scribner, K. T. (n.d.). Contiguity of landscape features pose barriers to gene flow among American marten (Martes americana) genetic clusters in the Upper Peninsula of Michigan. *Landscape Ecology*, *31*(5), 1051–1062. https://doi.org/10.1007/s10980-015-0331-0

Hutchinson, G.E. (1957) Concluding remarks. Cold Spring Harbor Symposium.

Sium. Quantitative biology. 22 415-427

Jetz, W., Wilcove, D. S., & Dobson, A. P. (2007). Projected Impacts of Climate and Land-Use Change on the Global Diversity of Birds. *PLoS Biology*, *5*(6), e157. https://doi.org/10.1371/journal.pbio.0050157

Keesing, F. (2000). Cryptic consumers and the ecology of an African savanna (vol 50, pg 205, 2000). *Bioscience*, *50*(6), 477.

Lidicker, W.Z. (1975) The role of dispersal in the demography of small mammals. Small Mamm 5:103–128

MacArthur, R.H. (1972) Geographical Ecology. Harper & Row, New York, NY.

Maclean, J. E., Goheen, J. R., Doak, D. F., Palmer, T. M., & Young, T. P. (2011). Cryptic herbivores mediate the strength and form of ungulate impacts on a long-lived savanna tree. *Ecology*, *92*(8), 1626–1636.

Manel, S., & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology & Evolution*, *28*(10), 614–621. https://doi.org/10.1016/j.tree.2013.05.012

Martinez-Gallardo, R., & Sanchez-Cordero, V. (1993). DIETARY VALUE OF FRUITS AND SEEDS TO SPINY POCKET MICE, HETEROMYS-DESMARESTIANUS (HETEROMYIDAE). *Journal of Mammalogy*, *74*(2), 436–442. https://doi.org/10.2307/1382400

Melbourne, B. A., & Hastings, A. (2008). Extinction risk depends strongly on factors contributing to stochasticity. *Nature*, *454*(7200), 100–103. https://doi.org/10.1038/nature06922

Moscarella, R. A., Hoffman, S. M. G., Myers, P., Yahnke, C. J., & Lundrigan, B. L. (2019). Genetic and demographic analysis of invasive Peromyscus leucopus in the northern Great Lakes region. *Journal of Mammalogy*, *100*(2), 345–353. https://doi.org/10.1093/jmammal/gyz053

Olden, J. D. (2006). Biotic homogenization: a new research agenda for conservation biogeography. *Journal of Biogeography*, *33*(12), 2027–2039. https://doi.org/10.1111/j.1365-2699.2006.01572.x

Orrock, J. L., Danielson, B. J., Burns, M. J., & Levey, D. J. (2003). *SPATIAL ECOLOGY OF PREDATOR-PREY INTERACTIONS: CORRIDORS AND PATCH SHAPE INFLUENCE SEED PREDATION*. *Ecology* (Vol. 84). Retrieved from https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/02-0439

Otero-Jiménez, B., Vandermeer, J. H., & Tucker, P. K. (2018). Effect of coffee agriculture management on the population structure of a forest dwelling rodent (Heteromys desmarestianus goldmani). *Conservation Genetics*, *19*(2), 495–499. https://doi.org/10.1007/s10592-017-1016-9

Pardini, R., Faria, D., Accacio, G. M., Laps, R. R., Mariano-Neto, E., Paciencia, M. L. B., … Baumgarten, J. (2009). The challenge of maintaining Atlantic forest biodiversity: A multi-taxa conservation assessment of specialist and generalist species in an agro-forestry mosaic in southern Bahia. *Biological Conservation*, *142*(6), 1178–1190. https://doi.org/10.1016/j.biocon.2009.02.010

Pardini, R. (2004). Effects of forest fragmentation on small mammals in an Atlantic Forest landscape. *Biodiversity and Conservation*, *13*(13), 2567–2586. https://doi.org/10.1023/B:BIOC.0000048452.18878.2d

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Perfecto, I., & Vandermeer, J. (2002). Quality of agroecological matrix in a tropical montane landscape: Ants in coffee plantations in southern Mexico. *Conservation Biology*, *16*(1), 174–182. https://doi.org/10.1046/j.1523-1739.2002.99536.x

Perfecto I, Vandermeer J, Wright A (2010). Nature's matrix: linking agriculture, conservation and food sovereignty. Routledge.

Peterman, W. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods Ecol Evol*, *9*, 1638–1647. https://doi.org/10.1111/2041-210X.12984

Peterman, W. E., Connette, G. M., Semlitsch, R. D., & Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Mol Ecol*, *23*(10), 2402–2413. https://doi.org/10.1111/mec.12747

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Scheibe, J. S. (1984). THE EFFECTS OF WEATHER, SEX AND SEASON ON THE NOCTURNAL ACTIVITY OF PEROMYSCUS-TRUEI (RODENTIA). *Southwestern Naturalist*, *29*(1), 1–5. https://doi.org/10.2307/3670763

Silva, M., Hartling, L., & Opps, S. B. (2005). Small mammals in agricultural landscapes of Prince Edward Island (Canada): effects of habitat characteristics at three different spatial scales. *Biological Conservation*, *126*(4), 556-568.

Smart, S. M., Thompson, K., Marrs, R. H., Le Duc, M. G., Maskell, L. C., & Firbank, L. G. (n.d.). Biotic homogenization and changes in species diversity across human-modified ecosystems. https://doi.org/10.1098/rspb.2006.3630

Tucker, M. A., Böhning-Gaese, K., Fagan, W. F., Fryxell, J. M., Van Moorter, B., Alberts, S. C., … Mueller, T. (2018). Moving in the Anthropocene: Global reductions in terrestrial

mammalian movements. *Science (New York, N.Y.)*, *359*(6374), 466–469. https://doi.org/10.1126/science.aam9712

Umetsu, F., & Pardini, R. (2007). Small mammals in a mosaic of forest remnants and anthropogenic habitats-evaluating matrix quality in an Atlantic forest landscape. *Landscape Ecology*, *22*(4), 517–530. https://doi.org/10.1007/s10980-006-9041-y

van Etten, J. (2017). *R* Package **gdistance** : Distances and Routes on Geographical Grids. *Journal of Statistical Software*, *76*(13). https://doi.org/10.18637/jss.v076.i13

Vázquez, E. & Reid, F. (2016). *Peromyscus gymnotis*. *The IUCN Red List of Threatened Species* 2016: e.T16666A22361063. http://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T16666A22361063.en. Downloaded on 2 June 2019.

White, A. J., Poulin, R. G., Wissel, B., Doucette, J. L., & Somers, C. M. (2012). Agricultural land use alters trophic status and population density of deer mice (*Peromyscus maniculatus*) on the North American Great Plains. *Canadian Journal of Zoology*, *90*(7), 868–874. https://doi.org/10.1139/z2012-055

Woinarski, J. C. Z., Armstrong, M., Brennan, K., Fisher, A., Griffiths, A. D., Hill, B., … Young, S. (2010). Monitoring indicates rapid and severe decline of native small mammals in Kakadu National Park, northern Australia. *Wildlife Research*, *37*(2), 116–126. https://doi.org/10.1071/wr09125

Young, H. S., McCauley, D. J., Dirzo, R., Goheen, J. R., Agwanda, B., Brook, C., … Helgen, K. M. (2015). Context-dependent effects of large-wildlife declines on small-mammal communities in central Kenya. *Ecological Applications*, *25*(2), 348–360. https://doi.org/10.1890/14-0995.1

Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology*, *27*(6), 777–797. https://doi.org/10.1007/s10980-012-9737-0

Chapter 3 Comparing The Population Structure Of A Forest Dwelling Rodent (*Heteromys desmarestianus goldmani*) In A Coffee Production Region And A Continuous Forest

1 Introduction

Tropical forests, the most biodiverse terrestrial ecosystems on the planet, are being transformed by human activities. The expansion and intensification of agriculture is one of the largest threats to these ecosystems (FAO 2016). This has led most tropical forests to exist as fragments embedded within a matrix of agricultural land uses (Perfecto et al. 2010). Organisms that previously existed within these continuous forests, now must learn to navigate through an agricultural environment in order to maintain their populations. For this reason, it is important to understand species' responses to these landscape changes in order to develop accurate conservation and management strategies. However, the influence of agricultural landscapes on tropical species is poorly understood and this lack knowledge limits our ability to implement successful strategies.

Rodents are an important part of tropical ecosystems but are sparsely represented in the scientific literature on conservation management. They represent 40% of all mammalian diversity and since 1992 more than 40 species and 12 new genera have been discovered (Amori and Gippoliti 2003). Additionally, they play important roles in food webs as prey and as predators of invertebrates, fungi, seeds and other vegetation (King 1985, Correa and Roa 2005). Rodents are also seed (Vieira et al. 2003, Galetti et al. 2015) and mycorrhizal fungi dispersers (Janos et al. 1995), and thus are important in promoting tree regeneration (Michel et al. 2007).

Many of the roles rodents play in ecosystems have been identified as critical and non-ecologically redundant (Amori and Gippoliti 2003, Andersen et al. 2018). For these reasons, it is important to understand the responses of rodent populations to human modified landscapes.

Despite their large role in ecosystem function, tropical rodents are not well studied, and conservation initiatives are biased to more charismatic mammal groups and species (Amori and Gippoliti 2003). This could be due to the fact that some rodent species benefit from human disturbance, either by the increase in resource availability or the elimination of predator species (Arce-Peña et al. 2019). However, available research has found that not all rodent species are equally resilient to disturbances, and many species can be impacted by habitat loss and fragmentation (Umetsu and Pardini 2007, Gibson et al. 2013, Woinarski et al. 2010). This is especially alarming since 40% of all rodent genera are monotypic, and phylogenetic diversity could be lost at a rapid rate (Amori and Gippoliti 2003). Additionally, rodents have the highest extinction rate among mammals, representing 50-52% of all extinctions in the past 500 years (Amori and Gippoliti 2003).

Our study aims to understand the effects of habitat fragmentation and agricultural management on the population structure of a common forest rodent *Heteromys desmarestianus goldmani*. Development of new and affordable molecular methods have allowed researchers to explore population effects on small or elusive species. We take advantage of these new methods to calculate the level of population connectivity, through measures of gene flow, for the species. We sampled *H. d. goldmani* in both a continuous forest and a coffee production region, composed of forest fragments embedded within a matrix of varying coffee production management practices. We compared the genetic structure of *H. d. goldmani* between the two

locations in order to gain a better understanding of the effects of coffee agriculture development and fragmentation on the species. We expect that *H. d. goldmani* individuals in the continuous forest will have stronger population connectivity, higher gene flow and lower population structure, when compared to individuals in the area of coffee production.

2 Materials and Methods

2.1 Study System and Sampling

2.1.1 Study sites

Coffee production in Latin America supplies most of the coffee consumed worldwide (FAO 2015). In this region coffee is produced in a variety of ways that follow an intensification gradient, ranging from rustic coffee production (e.g., coffee grown in the understory of tropical forest) to sun coffee (Moguel and Toledo 1999; Fig. 3-1b, d). It is common to find several of these management practices represented in small areas. For this reason, we chose to do our study in the tropical montane region of Soconusco in Chiapas, Mexico (Fig. 3-1b, d). This area is dominated by coffee production that varies in management intensity (Fig. 3-1). Farms included in the study have been producing coffee for 60 to 100 years. Although management practices vary over time, these farms have had similar management practices for at least the past 20 years (Perfecto and Vandermeer 2002).

El Triunfo Biosphere Reserve, in the state of Chiapas, Mexico, is one of the largest remaining contiguous forests in the country (Fig. 3-1a, b). El Triunfo Biosphere Reserve was the first to enroll in the *Man and the Biosphere* program from UNESCO in 1993. It is located in the Sierra Madre mountain range in southern Mexico and is considered the most diverse cloud forest in the country. The reserve has a total area of 119,177 ha, from which 20% is part of the nucleus

zone which is divided into 5 polygons distributed across the reserve and surrounded by a buffer zone that includes agricultural and residential land uses (Martinez-Melendez et al. 2009). Our research was conducted in the nucleus polygon I (Fig. 3-1b, c). This reserve is the closest continuous forest to the coffee production area included in the study and serves as a good comparison with respect to elevation, temperature, rainfall, and geographical region. The reserve is an ideal area to conduct research on the dynamics of populations in their natural environment.

2.1.2 Study species

Despite their important role in tropical ecosystems, little is known about rodent population structure and connectivity and the effect of human modified landscape on these species. In this study we worked with Goldman's spiny pocket mouse, *H. d. goldmani,* a rodent that inhabits moist forest habitats from southern Mexico to Panama (Fleming 1974). It is the most abundant forest rodent in these regions (Klinger 2007) and has a home range of 100 m^2 (Fleming 1974), which is small for a terrestrial rodent. Studies conducted in Costa Rica have shown that the diet of *H. desmarestianus* mainly consists of palm nuts and other seeds (Fleming 1983), making this rodent an important seed disperser in tropical regions (DeMattia et al. 2004).

2.1.3 Field Sampling

H. d. goldmani samples were collected from 2012 to 2017 in the coffee production study site and March-April 2018 in the forest reserve (i.e., El Triunfo Biosphere Reserve). In the coffee production sites, ear tissue samples were collected from six sites: three forest fragments and three coffee farms of various management levels that are adjacent to the forest fragments (Fig.3-1). To facilitate continuous animal collection across this complex landscape we followed roads and trails within the farms and the forest fragments, following the methodology of Otero Jiménez

et al. 2018. A similar method was followed at the forest reserve, where we collected ear tissue samples from mice collected along monitoring trails in the reserve. In both sites we sampled an area of approximately 4 km x 2 km (Fig. 3-1c, d),

Ear tissue samples were preserved in 20% DMSO buffer saturated with NaCl. Mice were captured using 22.9 cm x 7.6 cm x 8.9 cm Sherman live traps. Species, sex, life stage and GPS coordinates for each individual sample were recorded. Animals were handled in accordance with the University of Michigan's Committee on Use and Care of Animals.

2.2 Genetic Data

2.2.1 SNP library preparation

Genomic DNA was extracted from ear tissue samples using the Qiagen DNeasy blood and tissue kit following the manufacture's protocol (Qiagen Inc.) and eluted in water. DNA libraries were constructed using a triple enzyme RADseq (3RAD) protocol (Glenn et al. 2017). We digested approximately 100 ng of DNA from each individual with ClaI, BamHI, and MspI restriction enzymes. We ligated all fragments to internal adapters with indexing tags of 5-8 nucleotides. A reduced cycle PCR and the standard Illumina adapter ligation protocol using Kapa LTP library preparation kits (Kapa Biosystems, Wilmington, MA) were used to construct full-length libraries. Samples were indexed using iTru5 and iTru7 primers (Glenn et al. 2016) and pooled. We used Pippin Prep (Sage Science, Beverly, MA) to size-select for fragments between 473 bp and 578 bp. We sequenced 56 *H. d. goldmani* individuals from the coffee farms and 32 individuals from the forest reserve in separate (i.e., each site) 150bp pair-end Illumina HiSeq 4000 lanes at the University of Michigan DNA Sequencing Core.

2.2.2 SNP calling and filtering

First, we checked for adapter contamination and quality using FASTQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Next, we used the *process_radtags* script from Stacks 2.1 (Catchen et al. 2013) to identify reads from each individual in the sequence data and to trim all reads to 120 bp. Then we assembled *denovo* sequences for each study site (i.e., coffee production and forest reserve) following the 6 step Stacks 2.1 pipeline (Catchen et al. 2013): (1) *ustacks* builds the SNP loci for each individual, (2) *cstacks* then creates a catalog of loci for all samples, (3) *sstacks* matches loci from individuals to the catalog. All these steps are done for each read separately. (4) Next, the *tsv2bam* script was used to pull together the pair-end reads associated with each SNP and orient the data by locus. (5) The *gtsacks* program assembled and merged paired-end contigs, and called variant sites in each sample. ((6) Finally, we generated data sets for each species using the *population* program. These data sets contained only SNPs that were (1) found in >75% of the samples (r = 0.75); (2) had a minor allele frequency ≥ 5% (min_maf = 0.05); (3) had moderate heterozygosity ≤ 0.8; and (4) retained only a single SNP per RADtag (Catchen et al. 2013).

The generated datasets were further filtered for over-clustering using VCFtools (Danecek et al. 2011). We removed sites with a total depth > 2500 and retained sites that had a mean depth between 10 and 40. We verified that individuals were genotyped at ≥ 95 % of all loci. The remaining loci for each species were used for the subsequent analysis.

2.2.3 Population genomic analysis

2.2.3.1 Genetic Diversity

We calculated the genome wide diversity by measuring expected heterozygosity (H_E), observed heterozygosity (H_O), nucleotide diversity (π), and inbreeding coefficient (F_{IS}) for each population separately for the filtered SNP library using the *population* function in Stacks 2.1 (Catchen et al. 2013).

2.2.3.2 Isolation by distance

To assess the influence of geographic distance on the genetic structure we calculated isolation-by-distance (IBD). We estimated pairwise Euclidean and genetic distances in Genalex (Peakall and Smouse 2012). Next, we measured the correlation between pairwise genetic distance of all sampled individuals, and Euclidian distance using a simple Mantel test in R (R Core Team, 2018). In addition, we estimated a Mantel correlogram based on 100 m distance classes and calculated significance using Spearman correlation based on 10,000 permutations using the *ecodist* R package. These analyses were done for each study site.

2.2.3.3 Clustering

We conducted population structure analysis using the SNPs retained after filtering across all samples. Bayesian clustering analysis using both fastSTRUCTURE v1.0 (Raj et al. 2014) and TESS3 R package (Caye et al. 2016) for samples of each species separately. The admixture model of fastSTRUCTURE assumes that each individual has some proportion of ancestry originating from a number (K) of gene pools, which is reflected in the inferred ancestry coefficients (Q value). We also applied the logistic prior to the model which allows us to detect

more subtle signals of genetic structure. TESS3 acts under similar assumptions but the model incorporates the geographic distance between samples, and thus it is able to detect weaker population structure. In both cases, we tested values of K ranging between 1-10, with 20 replicate runs per K. For fastSTRUCTURE, optimal K was chosen using a combination of the 'chooseK' script, and cross validation error, and for TESS3, we used cross entropy scores to determine optimal K value. Individual ancestry coefficients were plotted using *distruct* for fastSTRUCTURE results and for the TESS3 results were plotted in R. We used geographic location information and the inferred ancestry coefficient for each individual for the best supported K to generate the population membership maps in ArcGIS. Individuals with $Q > 0.8$ for any given cluster were assigned to that cluster, all other individuals were considered admixed.

3 Results

3.1 Sequencing and genotyping

We obtained data from all *H. d. goldmani* samples, 32 from the forest reserve and 56 from the coffee production site. Sequencing yielded an average of 493 million reads for each site and 2416 and 9770, after filtering, for the forest and coffee production sites, respectively (Table 3-1), with a mean coverage of 30X (range = 10-40X).

3.2 Genetic Diversity

Results for genetic diversity measures show significant differences between sampling sites (Fig. 3-2, Table 3-1). We observed slightly higher nucleotide diversity, expected and observed heterozygosity for *H. d. goldmani* samples from the forest reserve than from those in

the coffee production (Fig. 3-2, Table 3-1). Measures for the inbreeding coefficient (F_{IS}) were similar between both sampling sites suggesting the same levels of inbreeding in both locations.

3.3 Isolation by distance

We did not find significant patterns of isolation-by-distance for *H. d. goldmani* in any of our study sites when including all samples. Mantel r values were not significant for the forest reserve (Mantel r = 0.065, p > 0.1; Table 3-1) or coffee production (Mantel r = 0.105, p > 0.1; Table 3-1). Results for the Mantel correlogram for IBD at different distance classes show a similar pattern for both of the study sites. We found significant positive correlations between genetic and geographic distance at shorter distance classes (0-300m) for both sampling sites (Fig. 3-1), with the correlation being slightly stronger in the forest reserve. This analysis also revealed some differences in the IBD patterns for *H. d. goldmani* between sites. For example, we found a significant positive correlation at the 900 m distance class that was not identified in the coffee production samples (Fig. 3-1).

3.4 Clustering

We analyzed population structure using two different Bayesian methods, fastSTRUCTURE (Raj et al. 2014) and TESS3 (Cave et al. 2012). Results for both methods yielded similar patterns of structure for *H. d. goldmani* from both sites. For the forest reserve samples fastSTRUCTURE results supported the presence of 1 to 6 clusters (K), while TESS3 showed support for K = 1 having the lowest cross-entropy value (Fig. 3-4). When plotting the individual cluster assignments for fastSTRUCTURE we noticed that K values from 3-6 had the presence of ghost clusters (i.e., no individuals assigned to these clusters). We could not identify any spatial pattern to individual assignments for K 2 to 6 (Fig. 3-5a), thus we interpret K=1 as

being the best supported number of clusters. These results suggest that there is a high degree of connectivity in *H. d. goldmani* across the forest reserve.

Results for *H. desmarestianus* samples from the coffee farms showed signals of subtle genetic structure. Results from fastSTRUCTURE supported the presence of 2 to 8 clusters (K), while TESS3 showed support for K of 1 or 2, with K = 2 having a slightly lower cross-entropy score than K = 1, representing a higher likelihood (Fig. 3-4). We plotted individual cluster membership for each of the potential best supported K values for each method (Fig. 3-5b, c). For fastSTRUCTURE we found the presence of ghost clusters in all K values except for K = 2 and K = 3 (Fig. 3-5b). Patterns of individual assignment are maintained through different K values, highlighting the differentiation of 2 clusters. These results overlap with the pattern found in the TESS K=2 individual assignment (Fig. 3-5c). For this reason, we interpreted K=2 as being the supported number of clusters for *H. d. goldmani* samples in the coffee farms. For both methods the first cluster is composed of all individuals in the high, medium, and low intensity coffee farms, and individuals from the middle forest fragment (Fig. 3-6). The second cluster is composed of individuals sampled in the eastern forest fragment at the border of the high intensity coffee farm (Fig. 3-5d). These results suggest subtle population structure in *H. d. goldmani* within the coffee agroecosystem.

4. Discussion

The aim of this study was to compare the population structure of the common forest rodent (*H. d. goldmani*) between a continuous forest (El Triunfo Biosphere Reserve) and a coffee agroecosystem. We used molecular and geographic data to infer patterns of population connectivity and structure. Our results showed subtle reduction in population connectivity of *H.*

d. goldmani in the coffee farms when compared to the continuous forest. Our results suggest the coffee agroecological landscape sampled is able to maintain *H. d. goldmani* populations and high degrees of genetic diversity.

Our results showed higher heterozygosity and nucleotide diversity for the *H. d. goldmani* forest population, when compared to individuals in the coffee farms. Although significant these differences represented only a slight difference, suggesting that the coffee agroecosystem is capable of supporting *Heteromys* migration and thus genetic diversity. Despite our best efforts to sample across the coffee production landscape, we did not capture *Heteromys* individuals in large regions of the high intensity coffee farm, especially on the east side of the study area. Based on these observations we believe that the success of *Heteromys* populations in this landscapes is due to the fact that the high intensity coffee production farm sampled is surrounded by forest fragments and lower intensity coffee production which could be serving as source populations, thus minimizing the negative effect of this intensified agricultural practice on *Heteromys*.

We found that there is a subtle signal of population structure in *H. d. goldmani* samples from the coffee farms. These results support our initial predictions where we expected *Heteromys* individuals in the coffee agroecosystem to show higher levels of genetic structure than those found in the forest reserve. The pattern of the genetic clustering suggest that the high intensity coffee production could be impeding movement of individuals on the eastern side of the coffee production site. However, we found similar levels of genetic diversity between the forest and coffee farm *Heteromys*. This suggests that the barriers to connectivity in the coffee agroecosystem are not enough to significantly impede migration.

Our results support other patterns found in the literature for *Heteromys desmarestianus*. Our literature search yielded a single study that explored the population structure of *Heteromys* in agricultural systems (Otero Jimenez et al. 2018). This study used microsatellite markers and found stronger patterns of genetic structure within the same coffee production region (Otero Jimenez et al. 2018). This stronger population structure could be due to the difference in genetic markers used, and that in this case, microsatellite markers are more informative than the SNP markers. A study exploring rodent diversity identified coffee agroecosystems as an important refuge for this group within the agricultural landscape (Caudill and Rice 2016). Caudill and Rice (2016) found *H. desmarestianus* in forest fragments and coffee farms of varying management and identified vegetation cover as an important factor in driving rodent diversity and density. These findings support the results observed in our study where *Heteromys* was present in all 3 types of farms and seemed to have limited connectivity in high intensity coffee farms with lower tree cover. Arce-Peña et al. (2019) explored the changes in population density of several rodent species, including *H. desmarestianus,* in forest fragments and found increases in density of *Heteromys* in the forest fragment but decreases in the buffer zones around the forest. These results indicate that *H. desmarestianus* is susceptible to landscape changes, and supports Caudill and Rice (2016) and our findings suggesting that coffee agroecosystems can support healthy populations of this species.

This study has generated novel data on the population structure of *Heteromys* species in its natural environment, a continuous tropical forest. Other studies had evaluated the genetic diversity and structure of *H. desmarestianus* in fragmented landscapes (Arce-Peña et al. 2019, Otero Jimenez et al. 2018), but the interpretation of the results from these studies was limited since there was no information on the structure of 'undisturbed' *H. desmarestianus* populations.

Results from this study provide baseline data on the structure of a forest rodent species in a continuous forest. Genetic data can help in the development of conservation and agricultural management, and *H. desmarestianus* could be used as a landscape connectivity indicator species since it is common in forested landscapes across its range. We hope that this study will encourage others to investigate the population structure of different species within their natural habitats in order to increase our understanding of environmental changes on populations.

5 Tables

Table 3-1 Number of samples, number of SNP markers, genetic diversity measures and IBD correlation for *H. d. goldmani* forest and coffee production samples. Values in parenthesis are the standard error.

	Study Site	
	Forest	Coffee Farms
Samples	32	56
SNPs	2416	9770
Observed Heterozygosity	0.277 (± 0.003)	0.247 (± 0.001)
Expected Heterozygosity	0.302 (± 0.003)	0.275 (± 0.001)
Inbreeding Coefficient	0.105 (± 0.03)	0.109 (± 0.05)
Nucleotide diversity	0.307 (± 0.003)	0.278 (± 0.001)
Isolation-by-Distance	0.065 (p = 0.156)	0.105 (p = 0.131)

6 Figures

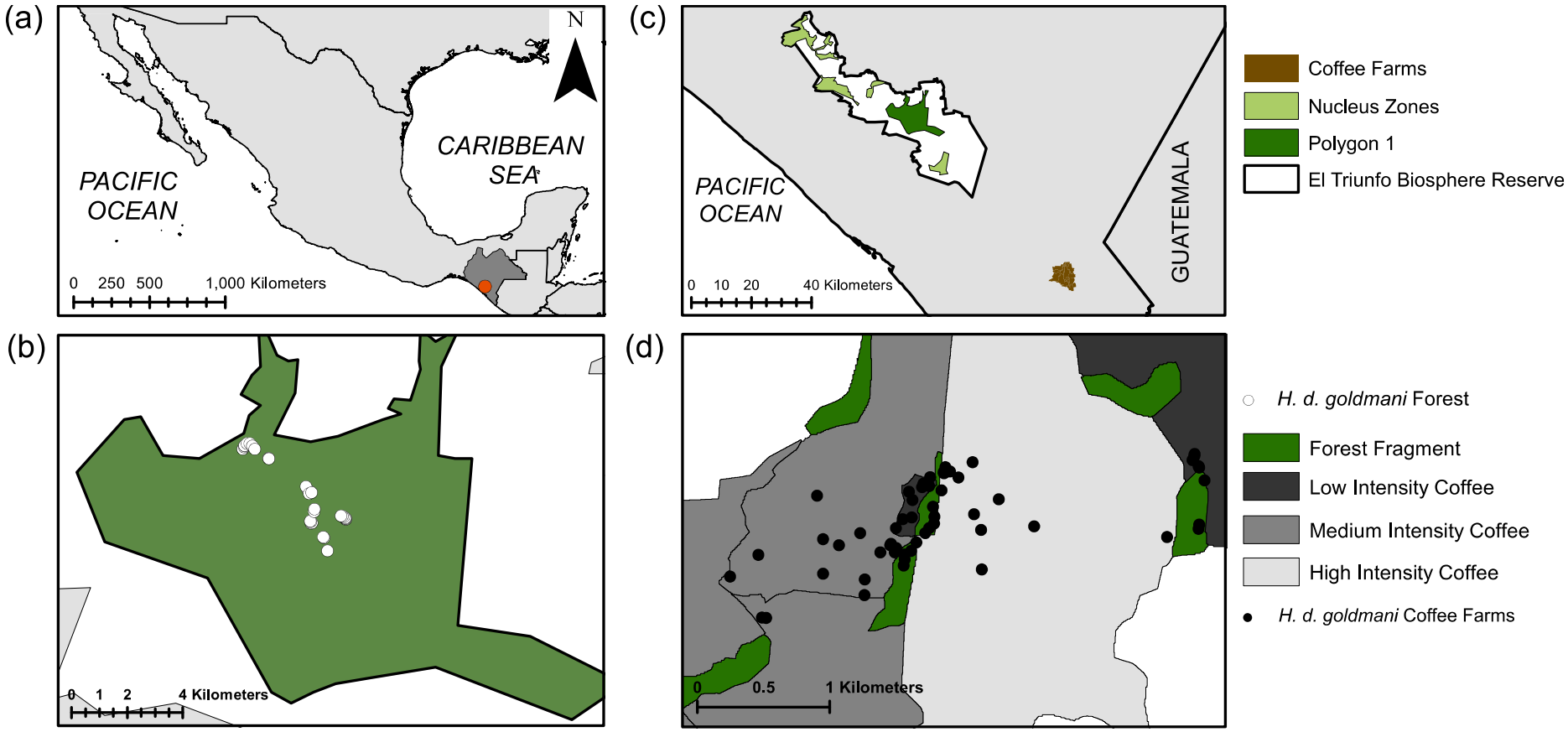


Figure 3-1 Map of sampling sites. (a) Location of Chiapas, Mexico, red dot represents to location of the study sites within the state. (b) Study sites: El Triunfo Biosphere reserve and the coffee production site (tan). (c) Forest reserve sampling site (Nucleus Zone 1). (d) Coffee production region.

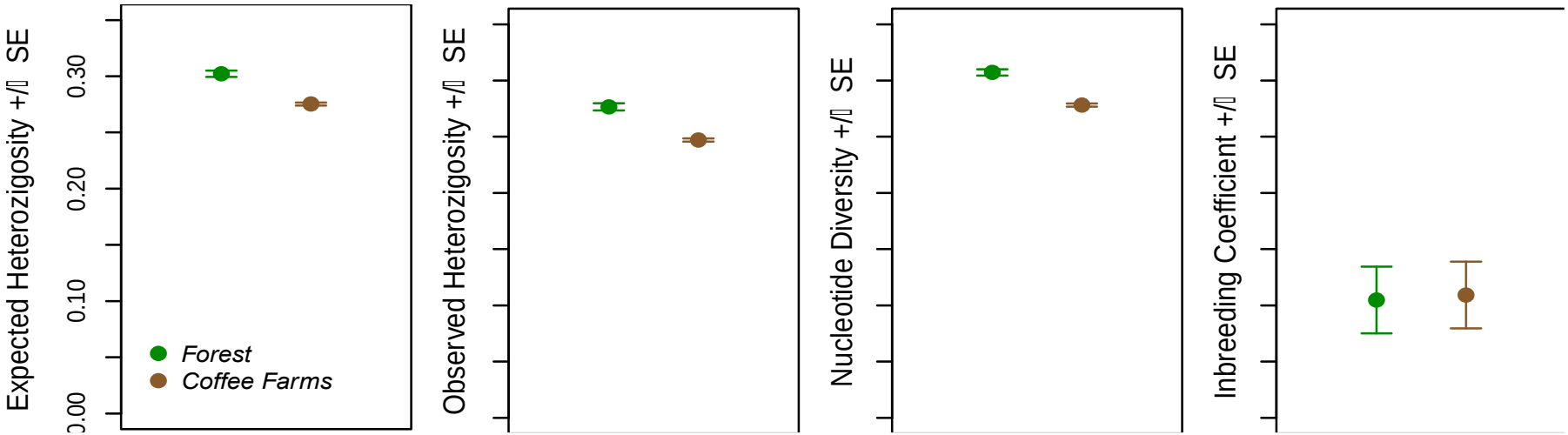


Figure 3-2 Mean values of genetic diversity measures and their standard error values for *Heteromys* forest reserve samples (green) and coffee production samples (tan).

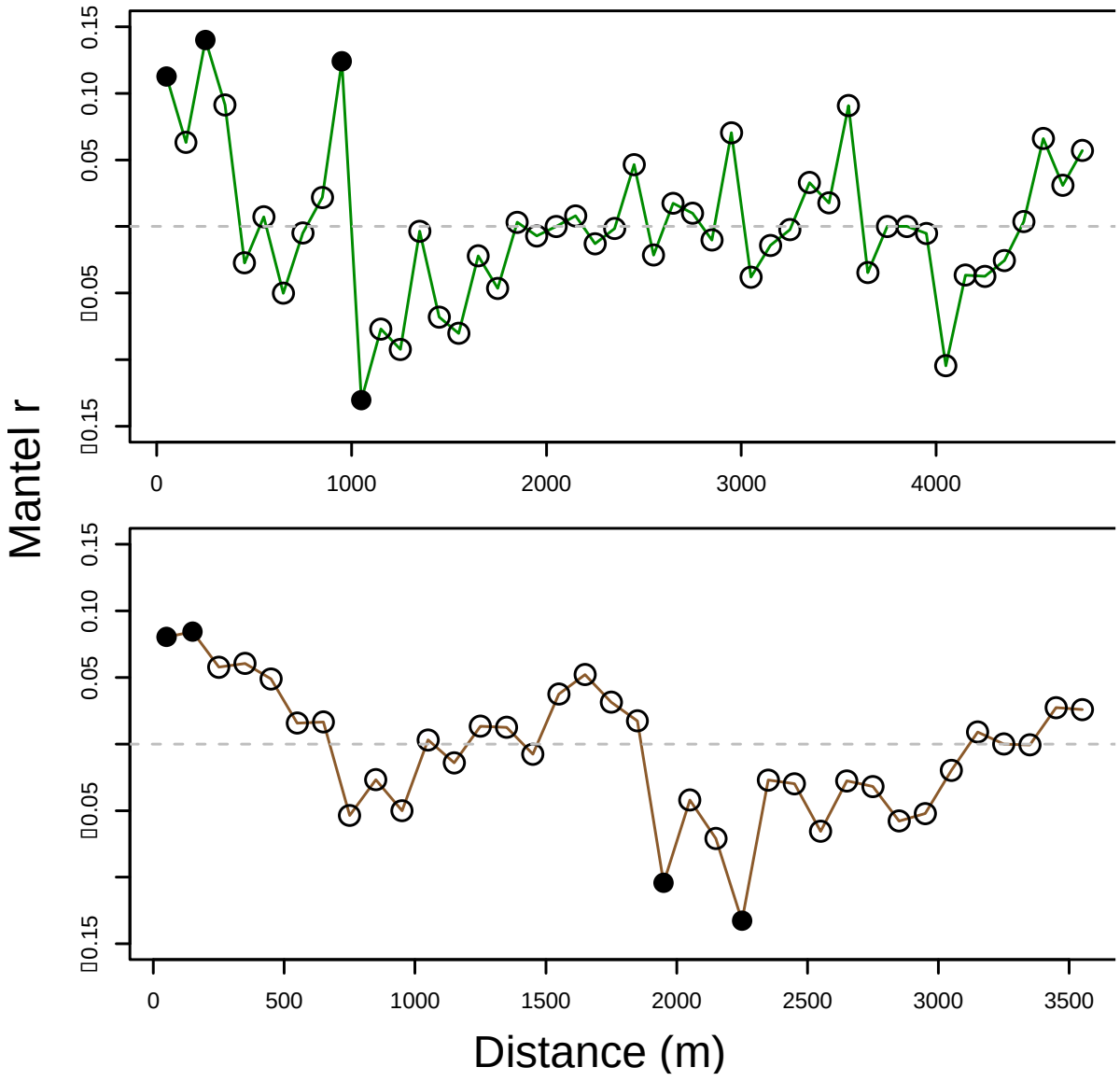


Figure 3-3 Isolation-by-distance Mantel correlogram results for *H. d. goldmani* (green, top) forest and (tan) coffee production samples. Filled black dots represent statistically significant correlations (p = < 0.05) between genetic and geographic distance.

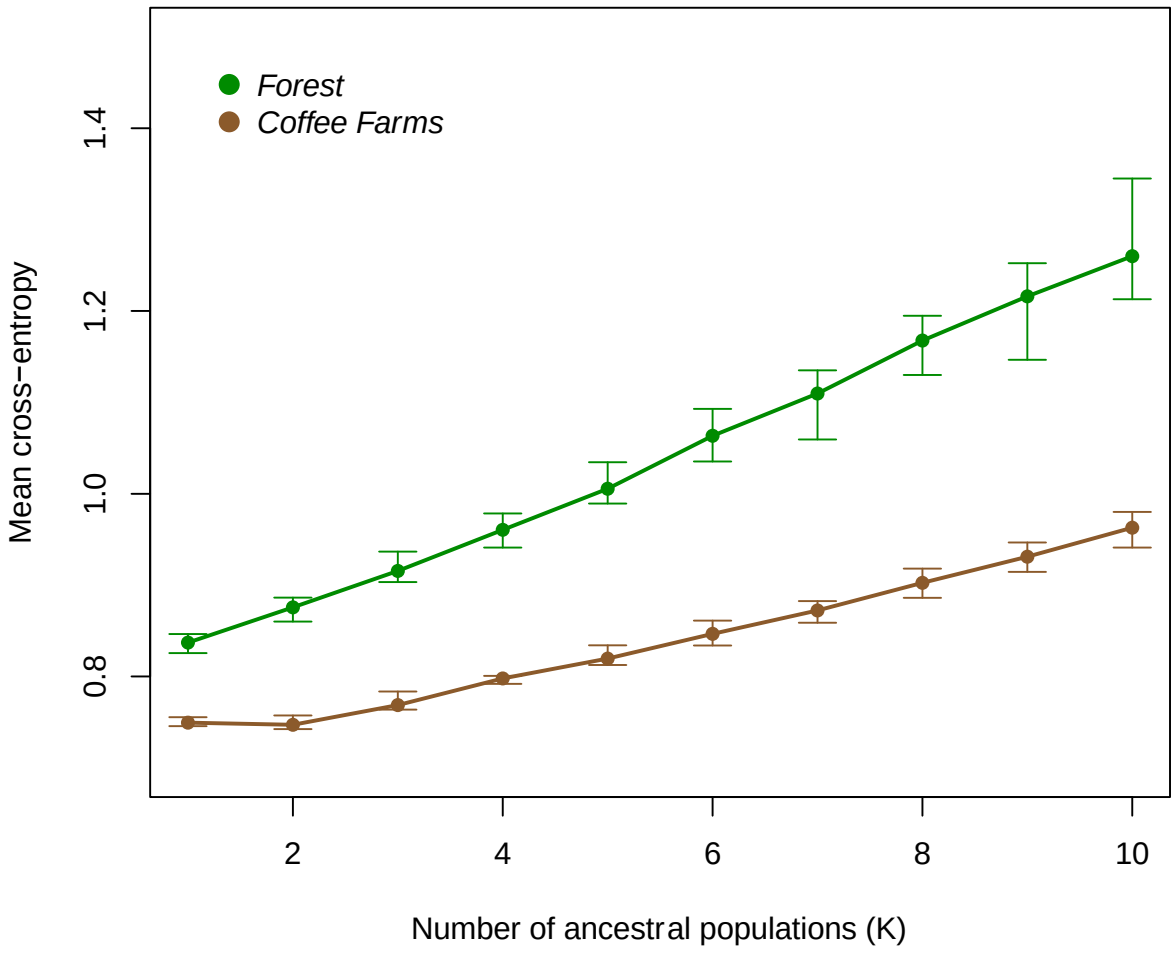


Figure 3-4 Mean cross-entropy results for TESS across all 10 replicates for each cluster (K) value for *H. d. goldmani* in the forest reserve and coffee farms.

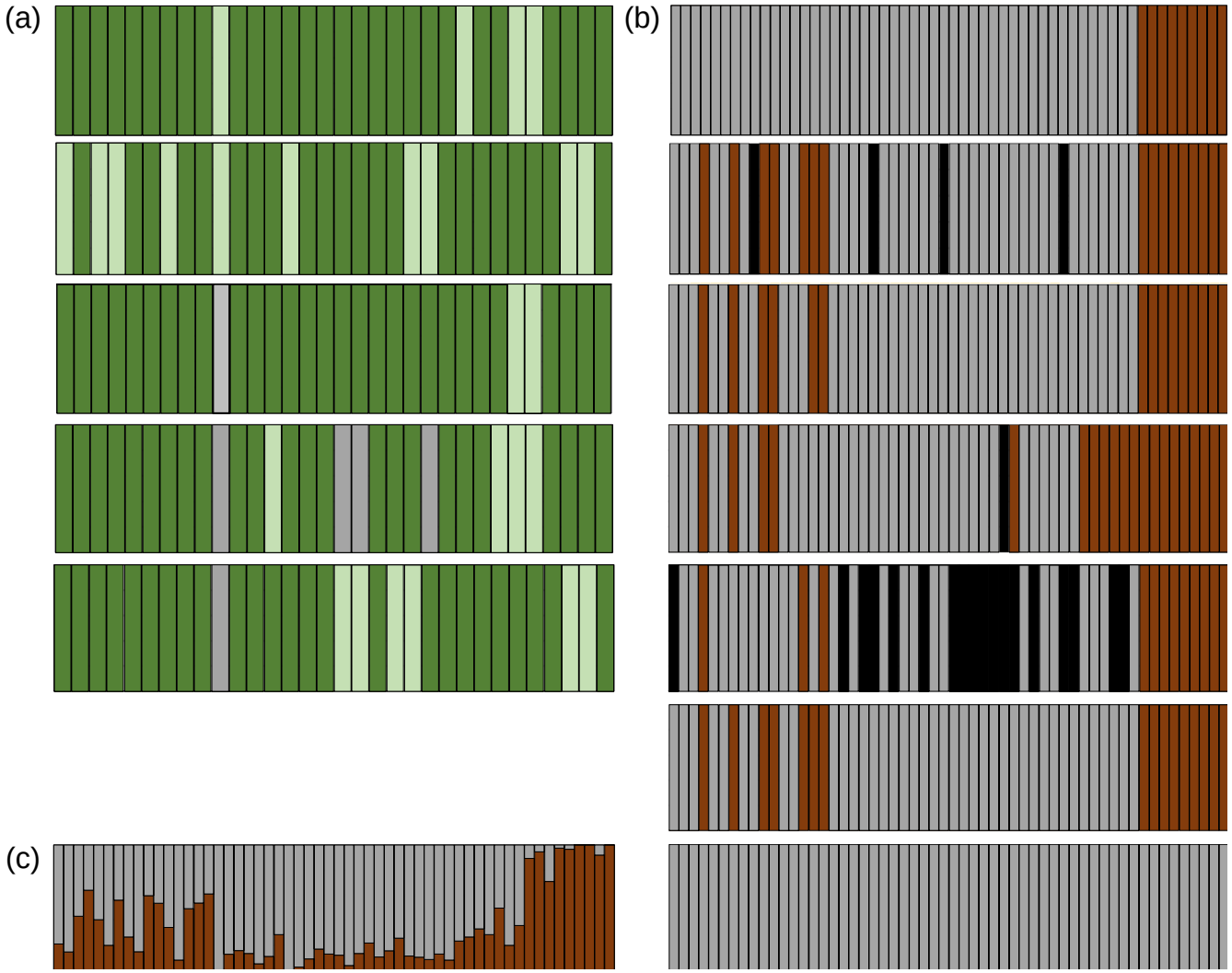


Figure 3-5 Genetic clustering results for *H.d goldmani* samples from the forest and coffee farms. Each column represents an individual and the colors represent the proportion of membership to a particular cluster. (a) fastSTRUCTURE results for K = 2 to 6 for the for the forest samples. Samples are arranged from north (left) to south (right) (see Fig.3-1c). (b) fastSTRUCTURE results for K = 2 to 8, and (c) TESS results for K=2 for coffee farm samples. Samples are arranged from west (left) to east (right) (see Fig. 3-1d).

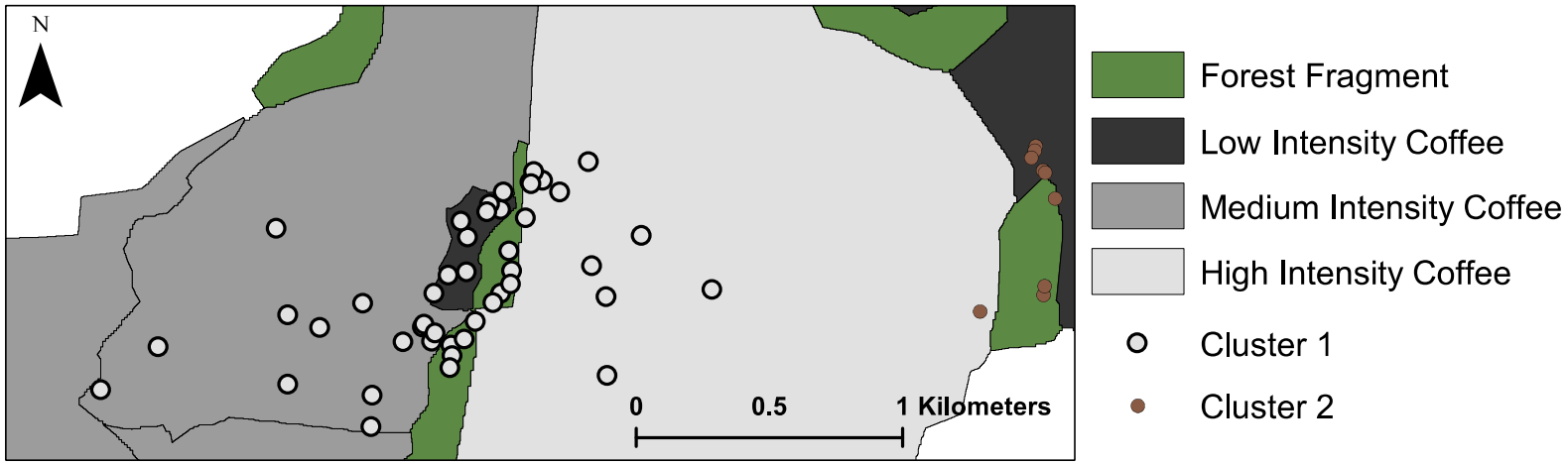


Figure 3-6 Map showing cluster (K) membership of *H. d. goldmani* individuals sampled in the coffee farms, based on fastSTRUCTURE K=2 individual assignment results.

References

Amori, G., & Gippoliti, S. (2003). A higher-taxon approach to rodent conservation priorities for the 21st century. *Animal Biodiversity and Conservation*, *26*(2), 1–18. https://doi.org/57406

Arce-Peña, N. P., Arroyo-Rodríguez, V., San-José, M., Jiménez-González, D., Franch-Pardo, I., Andresen, E., & Ávila-Cabadilla, L. D. (2019). Landscape predictors of rodent dynamics in fragmented rainforests. *Biodiversity and Conservation*, *28*(3), 655–669. https://doi.org/10.1007/s10531-018-1682-z

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Mol Ecol*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Caudill, S. A., & Rice, R. A. (2016). Do Bird Friendly® Coffee Criteria Benefit Mammals? Assessment of Mammal Diversity in Chiapas, Mexico. *PLOS ONE*, *11*(11), e0165662. https://doi.org/10.1371/journal.pone.0165662

Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. (2016). TESS3: fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources*, *16*(2), 540–548. https://doi.org/10.1111/1755-0998.12471

Correa, P. & Roa, A. (2005) Relaciones tróficas entre Oncifelis guigna, Lycalopex culpaeus, Lycalopex griseus y Tyto alba en un ambiente fragmentado de la zona central de Chile. Mastozool. Neotrop. 12: 57–60.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

DeMattia, E. A., Curran, L. M., & Rathcke, B. J. (2004). Effects of small rodents and large mammals on Neotropical seeds. *Ecology*, *85*(8), 2161–2170. https://doi.org/10.1890/03-0254

FAO (2015) Statistical Pocketbook Coffee 2015. FAO, Rome

FAO (2016) State of the World's Forests 2016. Forests and agriculture: land-use challenges and opportuni- ties. FAO, Rome

Fleming, T.H. (1974) The Population Ecology of Two Species of Costa Rican Heteromyid Rodents. Ecology 55(3):493-510

Fleming, T.H. (1983). Heteromys desmarestianus. *Costa Rican natural history (DH Janzen, ed.). University of Chicago Press, Chicago, Illinois*, 474-475.

Galetti, M., Bovendorp, R. S., & Guevara, R. (2015). Defaunation of large mammals leads to an increase in seed predation in the Atlantic forests. *Global Ecology and Conservation*, *3*, 824–830. https://doi.org/10.1016/j.gecco.2015.04.008

Gibson, L., Lynam, A. J., Bradshaw, C. J. A., He, F. L., Bickford, D. P., Woodruff, D. S., … Laurance, W. F. (2013). Near-Complete Extinction of Native Small Mammal Fauna 25 Years After Forest Fragmentation. *Science*, *341*(6153), 1508–1510. https://doi.org/10.1126/science.1240495

Glenn, T., Bayona-Vásquez, N., Kieran, T., Pierson, T., Hoffberg, S., Scott, P., … Faircloth, B. (2017). Adapterama III: Quadruple-indexed, triple-enzyme RADseq libraries for about $1USD per Sample (3RAD). *BioRxiv*, 205799. https://doi.org/10.1101/205799

Glenn, T. C., Nilsen, R. A., Kieran, T. J., Sanders, J. G., Bayona-Vásquez, N. J., Finger, J. W., … Faircloth, B. C. (2019). Adapterama I: Universal stubs and primers for 384 unique

dual-indexed or 147,456 combinatorially-indexed Illumina libraries (iTru & iNext). *BioRxiv*, 049114. https://doi.org/10.1101/049114

Janos, D. P., Sahley, C. T., & Emmons, L. H. (1995). Rodent Dispersal of Vesicular-Arbuscular Mycorrhizal Fungi in Amazonian Peru. *Ecology*, *76*(6), 1852–1858. https://doi.org/10.2307/1940717

King, C.M. (1985). Interactions between woodland rodents and their predators. *The Ecology of Woodland Rodents : Bank Voles and Wood Mice*, 377–411. Retrieved from https://ci-nii-ac-jp.proxy.lib.umich.edu/naid/10027360558/

Klinger, R. (2007). Catastrophes, disturbances and density-dependence: population dynamics of the spiny pocket mouse (Heteromys desmarestianus) in a neotropical lowland forest. *Journal of Tropical Ecology*, *23*, 507–518. https://doi.org/10.1017/s0266467407004415

Martínez-Meléndez, N., Pérez-Farrera, M., & MartÍnez-Camilo, R. (2009). The Vascular Epiphyte Flora of El Triunfo Biosphere Reserve, Chiapas, México. *Rhodora*, *111*(948), 503–535. https://doi.org/10.3119/08-20.1

Michel, N., Burel, F., Legendre, P., & Butet, A. (2007). Role of habitat and landscape in structuring small mammal assemblages in hedgerow networks of contrasted farming landscapes in Brittany, France. *Landscape Ecology*, *22*(8), 1241–1253. https://doi.org/10.1007/s10980-007-9103-9

Moguel, P., & Toledo, V. M. (1999). Biodiversity conservation in traditional coffee systems of Mexico. *Conservation Biology*, *13*(1), 11–21. https://doi.org/10.1046/j.1523-1739.1999.97153.x

Otero-Jiménez, B., Vandermeer, J. H., & Tucker, P. K. (2018). Effect of coffee agriculture management on the population structure of a forest dwelling rodent (Heteromys

desmarestianus goldmani). *Conservation Genetics*, *19*(2), 495–499. https://doi.org/10.1007/s10592-017-1016-9

Perfecto, I., & Vandermeer, J. (2002). Quality of agroecological matrix in a tropical montane landscape: Ants in coffee plantations in southern Mexico. *Conservation Biology*, *16*(1), 174–182. https://doi.org/10.1046/j.1523-1739.2002.99536.x

Perfecto I, Vandermeer J, Wright A (2010). Nature's matrix: linking agriculture, conservation and food sovereignty. Routledge.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, *197*(2), 573–589. https://doi.org/10.1534/genetics.114.164350

Umetsu, F., & Pardini, R. (2007). Small mammals in a mosaic of forest remnants and anthropogenic habitats-evaluating matrix quality in an Atlantic forest landscape. *Landscape Ecology*, *22*(4), 517–530. https://doi.org/10.1007/s10980-006-9041-y

Vieira, E. M., Pizo, M. A., & Izar, P. (2003). Fruit and seed exploitation by small rodents of the Brazilian Atlantic forest. *Mammalia*, *67*(4), 533–540. https://doi.org/10.1515/mamm-2003-0407

Woinarski, J. C. Z., Armstrong, M., Brennan, K., Fisher, A., Griffiths, A. D., Hill, B., … Young, S. (2010). Monitoring indicates rapid and severe decline of native small mammals in Kakadu National Park, northern Australia. *Wildlife Research*, *37*(2), 116–126. https://doi.org/10.1071/wr09125

Appendix

Appendix 1

Tables

Appendix 1 1 Genetic diversity of *H. d. goldmani* by locus. Na = Number of Alleles; Ne = Number of effective alleles; Ho = observed heterozygosity; uHe = unbiased Heterozygosity; FIS = Fixation index

	HET-1	**HET-4**	**HET-27**	**HET-23**	**HET-32**	**HET-34**	**HET-37**	**HET-41**	**HET-42**	**HET-46**	**HET-57**
N	162	164	165	163	164	163	164	165	161	157	164
Na	14	8	11	7	6	11	9	5	7	11	12
Ne	5.426	4.394	3.835	2.925	3.465	8.065	4.246	2.656	3.840	7.566	7.057
Ho	0.667	0.640	0.642	0.644	0.671	0.791	0.811	0.309	0.764	0.707	0.799
uHe	0.818	0.775	0.741	0.660	0.714	0.879	0.767	0.625	0.742	0.871	0.861
Fis	0.183	0.171	0.131	0.021	0.057	0.097	-0.061	0.504	-0.033	0.185	0.069

Appendix 1 2 Relatedness (Lynch & Ritland, 1999) estimator between *Heteromys* samples

Number of pairs	13530
Sum	-40.785
Mean	-0.003
Median	-0.013
Standard Deviation	0.059
Standard Error	0.001
Min	-0.174
Max	0.478

Appendix 1 3 Geneland results for optimal number of genetic clusters (K) of *H. d. goldmani*.

Run	Number of Populations	Percentage (%)	Average Log (P)
19	6.0	31.4	-2599
16	6.0	30.0	-2618
8	6.0	30.0	-2622
12	6.0	32.4	-2636
17	6.0	31.6	-2643
18	6.0	32.2	-2649
7	6.0	30.1	-2660
4	6.0	30.0	-2660
13	6.0	34.7	-2661
6	6.0	30.5	-2665
10	6.0	32.1	-2666
20	6.0	33.8	-2668
11	6.0	30.0	-2674
1	6.0	31.4	-2680
14	6.0	34.1	-2684
2	6.0	31.2	-2688
9	6.0	32.3	-2691
15	6.0	32.9	-2694
5	6.0	33.1	-2700
3	6.0	32.5	-2716

Appendix 1 4 Mantel correlogram r and p-values for geographic (Euclidian) and genetic (ar) distances calculated at 100m distance classes for *H. d. goldmani*. Asterisk identify distance classes with significant correlations (P < 0.05).

Distance (m)	Mantel r	Significance (p-value <0.05)
100	0.078	*
200	0.063	*
300	0.041	*
400	0.026	
500	0.028	*
600	-0.008	
700	-0.019	
800	-0.027	
900	-0.033	*
1000	-0.039	*
1100	-0.008	
1200	-0.021	
1300	-0.001	
1400	-0.008	
1500	-0.003	
1600	-0.019	
1700	0.010	
1800	0.007	
1900	-0.017	
2000	-0.057	*
2100	-0.041	*
2200	-0.009	
2300	-0.026	
2400	-0.010	
2500	-0.009	
2600	-0.019	
2700	0.000	
2800	-0.020	

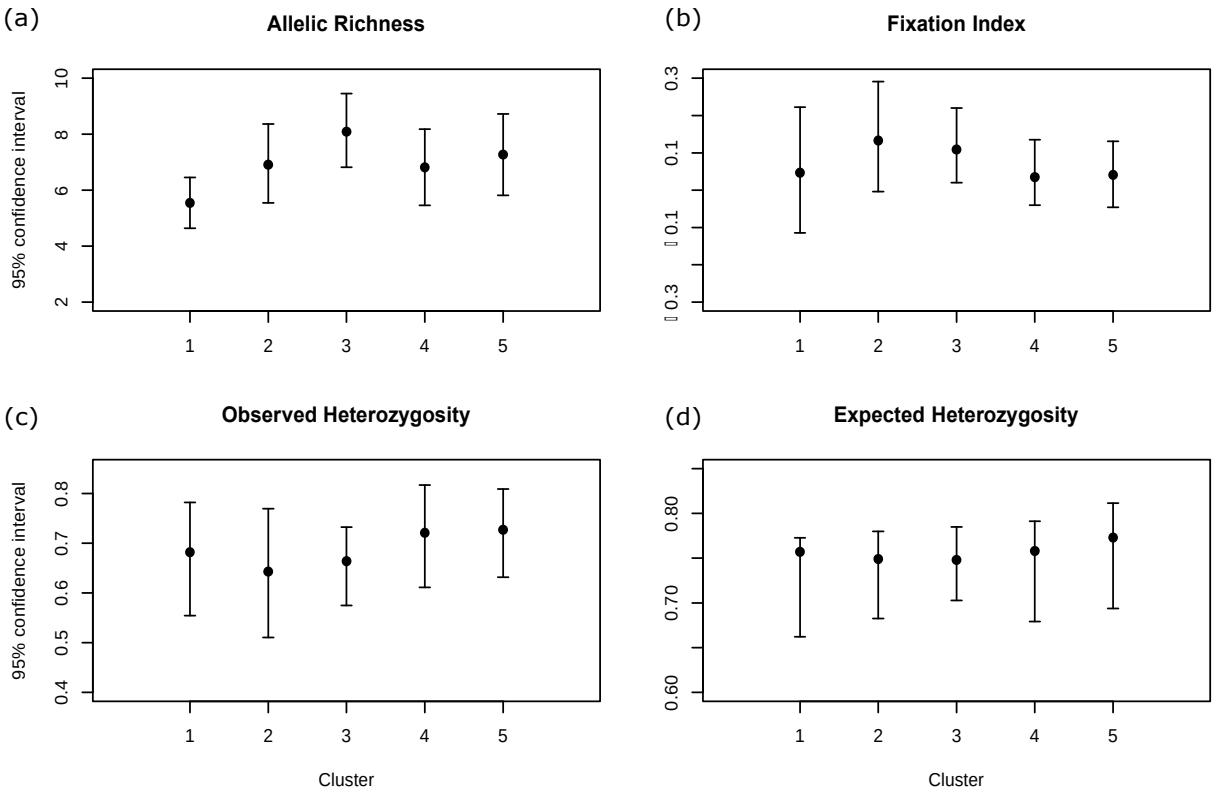
2900	-0.040	*
3000	-0.014	
3100	-0.008	
3200	0.000	
3300	0.000	
3400	-0.006	
3500	0.005	

Appendix 1 5 Spearman correlation results for all pairs of resistance surfaces and strength of correlation. Slope (S), tree cover (TC), elevation (E), riparian effect (RE), and streams (STR).

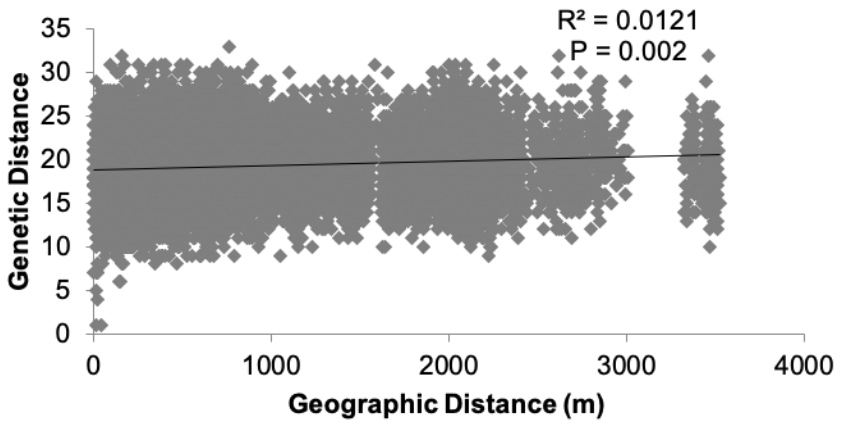
Layers	Spearman rs	p	Correlation
S + TC	0.491	2.20E-16	Moderate
S + STR	0.533	2.20E-16	Moderate
S + RE	0.597	2.20E-16	Moderate/ Strong
S + E	0.603	2.20E-16	Strong
TC + STR	0.730	2.20E-16	Strong
E + STR	0.817	2.20E-16	Strong
TC + RE	0.821	2.20E-16	Strong
TC + E	0.858	2.20E-16	Strong
STR + RE	0.887	2.20E-16	Strong
E + RE	0.908	2.20E-16	Strong

Figures

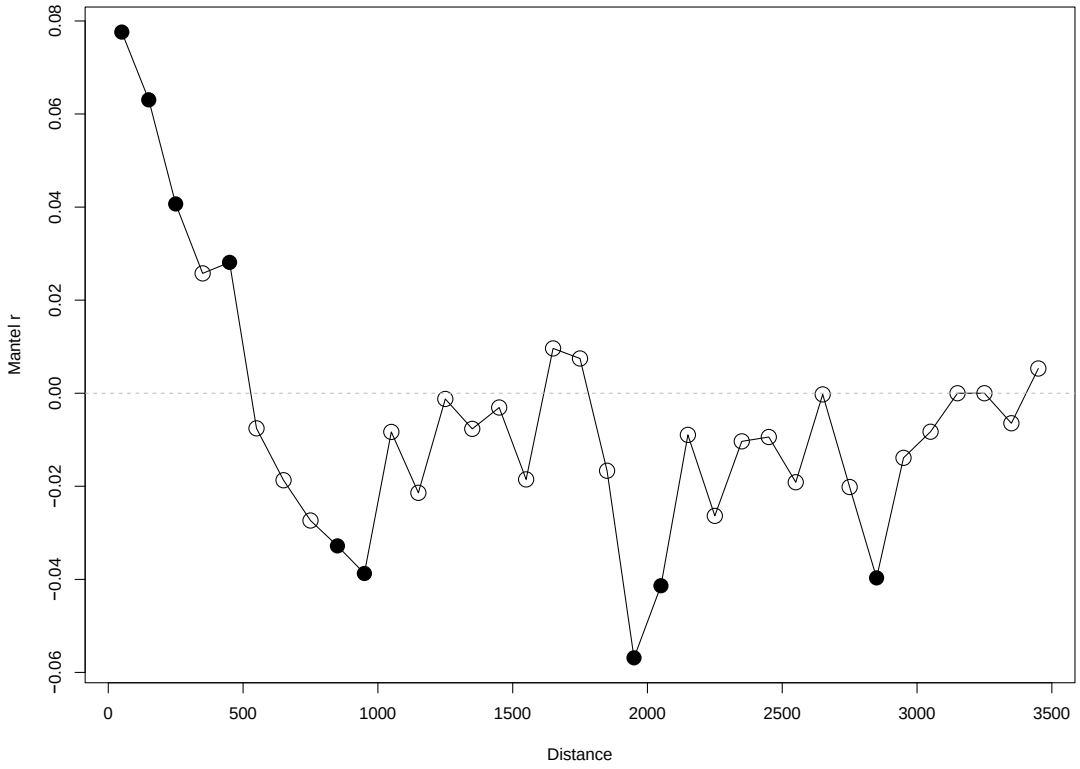
Appendix 1 6 Genetic diversity measures per genetic cluster identified by Geneland. Bars represent the 95% confidence interval.



Appendix 1 7 Isolation by distance for all samples of *H. d. goldmani*. Each diamond represents a pair of individuals.

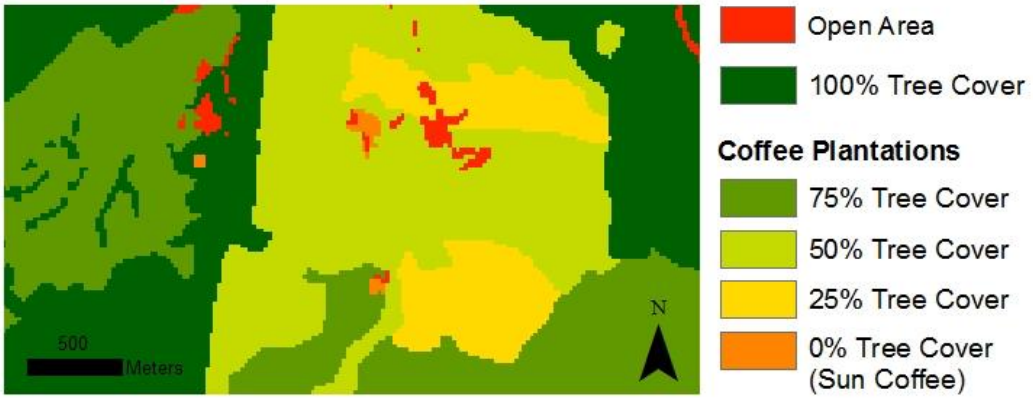


Appendix 1 8 Mantel correlogram for geographic (Euclidian) and genetic (ar) distances calculated at 100 m distance classes for *H. d. goldmani*. Significant values ($p < 0.05$) are shown in black. Their corresponding Mantel r values are in Table A 1-4.



Appendix 1 9 Raster map surfaces for tree cover: a) original input surface for ResistanceGA, and b) resulting ResistanceGA surface for the best model of tree cover based on genetic distance of *Heteromys* samples.

(a)



(b)